

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2003, 13:07:05 ; Search time 20 Seconds  
(without alignments)  
1840.522 Million cell updates/sec

Title: US-09-970-287-2  
Perfect score: 4751  
Sequence: 1 MGPPSLVLLLSATVFSLLG.....PEMKRPSKSLGQLWEGWEG 870

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgm2.6/prodata/1/iaa/5A\_COMB.pep: \*  
2: /cgm2.6/prodata/1/iaa/5B\_COMB.pep: \*  
3: /cgm2.6/prodata/1/iaa/6A\_COMB.pep: \*  
4: /cgm2.6/prodata/1/iaa/6B\_COMB.pep: \*  
5: /cgm2.6/prodata/1/iaa/PCUS\_COMB.pep: \*  
6: /cgm2.6/prodata/1/iaa/backfiles1.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4719.5	99.3	867	4	US-09-996-243-84
2	3776	79.5	650	4	US-09-886-319A-36
3	3175.5	66.8	871	4	US-09-773-426A-3
4	3117.5	65.6	867	4	US-09-668-673B-2
5	3059	64.4	1611	4	US-09-668-673B-16
6	2093	44.1	455	4	US-09-668-673B-14
7	1702	35.8	470	4	US-09-886-319A-35
8	1606.5	33.8	360	4	US-09-810-347-4
9	1590.5	33.5	372	4	US-09-810-347-2
10	1589	33.4	709	4	US-09-668-673B-3
11	1490	31.4	307	4	US-09-810-347-5
12	1183	24.9	309	4	US-09-810-347-6
13	827.5	17.4	510	1	US-08-484-493-11
14	827.5	17.4	510	1	US-08-484-494-11
15	827.5	17.4	510	2	US-08-345-212-11
16	827.5	17.4	510	3	US-09-249-003-11
17	827.5	17.4	510	4	US-09-685-844-11
18	565	11.9	160	4	US-09-668-673B-7
19	507	10.7	98	4	US-09-668-673B-22
20	405	8.5	96	4	US-09-668-673B-20
21	399.5	8.4	196	4	US-09-668-673B-18
22	370.5	7.8	520	4	US-09-773-426A-10
23	325	6.8	552	4	US-09-773-426A-9
24	286.5	6.0	406	4	US-09-227-357-187
25	277.5	5.8	556	1	US-08-445-586-2
26	276.5	5.8	554	1	US-08-445-586-7
27	238.5	5.0	583	1	US-08-445-586-9

28	238.5	5.0	583	4	US-09-646-683-1
29	234.5	4.9	550	1	US-08-484-493-2
30	234.5	4.9	550	1	US-08-484-494-2
31	234.5	4.9	550	2	US-08-345-212-2
32	234.5	4.9	550	3	US-09-249-003-2
33	234.5	4.9	550	4	US-09-685-844-2
34	232.5	4.9	583	1	US-08-484-493-14
35	232.5	4.9	583	1	US-08-484-494-14
36	232.5	4.9	583	2	US-08-345-212-14
37	232.5	4.9	583	3	US-09-249-003-14
38	232.5	4.9	583	4	US-09-685-844-14
39	230	4.8	547	4	US-09-252-991A-25705
40	215	4.5	551	1	US-08-484-493-15
41	215	4.5	551	1	US-08-484-494-15
42	215	4.5	551	2	US-08-345-212-15
43	215	4.5	551	3	US-09-249-003-15
44	215	4.5	551	4	US-09-685-844-15
45	208.5	4.4	533	1	US-08-445-586-10

ALIGNMENTS

RESULT 1

US-09-996-243-84  
; Sequence 84, Application US/09996243  
; Patent No. 6478825  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tamas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2730P1C13  
; CURRENT APPLICATION NUMBER: US/09/996,243  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28

Sequence 1, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 14, Appli  
Sequence 14, Appli  
Sequence 14, Appli  
Sequence 14, Appli  
Sequence 14, Appli  
Sequence 25705, A  
Sequence 15, Appli  
Sequence 15, Appli  
Sequence 15, Appli  
Sequence 15, Appli  
Sequence 10, Appli

1 PRIOR APPLICATION NUMBER: 60/084600  
2 PRIOR FILING DATE: 1998-05-07  
3 PRIOR APPLICATION NUMBER: 60/087106  
4 PRIOR FILING DATE: 1998-05-28  
5 PRIOR APPLICATION NUMBER: 60/087607  
6 PRIOR FILING DATE: 1998-06-02  
7 PRIOR APPLICATION NUMBER: 60/087609  
8 PRIOR FILING DATE: 1998-06-02  
9 PRIOR APPLICATION NUMBER: 60/087759  
10 PRIOR FILING DATE: 1998-06-02  
11 PRIOR APPLICATION NUMBER: 60/087827  
12 PRIOR FILING DATE: 1998-06-03  
13 PRIOR APPLICATION NUMBER: 60/088021  
14 PRIOR FILING DATE: 1998-06-04  
15 PRIOR APPLICATION NUMBER: 60/088025  
16 PRIOR FILING DATE: 1998-06-04  
17 PRIOR APPLICATION NUMBER: 60/088026  
18 PRIOR FILING DATE: 1998-06-04  
19 PRIOR APPLICATION NUMBER: 60/088028  
20 PRIOR FILING DATE: 1998-06-04  
21 PRIOR APPLICATION NUMBER: 60/088029  
22 PRIOR FILING DATE: 1998-06-04  
23 PRIOR APPLICATION NUMBER: 60/088030  
24 PRIOR FILING DATE: 1998-06-04  
25 PRIOR APPLICATION NUMBER: 60/088033  
26 PRIOR FILING DATE: 1998-06-04  
27 PRIOR APPLICATION NUMBER: 60/088326  
28 PRIOR FILING DATE: 1998-06-04  
29 PRIOR APPLICATION NUMBER: 60/088167  
30 PRIOR FILING DATE: 1998-06-05  
31 PRIOR APPLICATION NUMBER: 60/088202  
32 PRIOR FILING DATE: 1998-06-05  
33 PRIOR APPLICATION NUMBER: 60/088212  
34 PRIOR FILING DATE: 1998-06-05  
35 PRIOR APPLICATION NUMBER: 60/088217  
36 PRIOR FILING DATE: 1998-06-05  
37 PRIOR APPLICATION NUMBER: 60/088655  
38 PRIOR FILING DATE: 1998-06-09  
39 PRIOR APPLICATION NUMBER: 60/088734  
40 PRIOR FILING DATE: 1998-06-10  
41 PRIOR APPLICATION NUMBER: 60/088738  
42 PRIOR FILING DATE: 1998-06-10  
43 PRIOR APPLICATION NUMBER: 60/088742  
44 PRIOR FILING DATE: 1998-06-10  
45 PRIOR APPLICATION NUMBER: 60/088810  
46 PRIOR FILING DATE: 1998-06-10  
47 PRIOR APPLICATION NUMBER: 60/088824  
48 PRIOR FILING DATE: 1998-06-10  
49 PRIOR APPLICATION NUMBER: 60/088826  
50 PRIOR FILING DATE: 1998-06-10  
51 PRIOR APPLICATION NUMBER: 60/088858  
52 PRIOR FILING DATE: 1998-06-11  
53 PRIOR APPLICATION NUMBER: 60/088861  
54 PRIOR FILING DATE: 1998-06-11  
55 PRIOR APPLICATION NUMBER: 60/088876  
56 PRIOR FILING DATE: 1998-06-11  
57 PRIOR APPLICATION NUMBER: 60/089105  
58 PRIOR FILING DATE: 1998-06-12  
59 PRIOR APPLICATION NUMBER: 60/089440  
60 PRIOR FILING DATE: 1998-06-16  
61 PRIOR APPLICATION NUMBER: 60/089512  
62 PRIOR FILING DATE: 1998-06-16  
63 PRIOR APPLICATION NUMBER: 60/089514  
64 PRIOR FILING DATE: 1998-06-16  
65 PRIOR APPLICATION NUMBER: 60/089532  
66 PRIOR FILING DATE: 1998-06-17  
67 PRIOR APPLICATION NUMBER: 60/089538  
68 PRIOR FILING DATE: 1998-06-17  
69 PRIOR APPLICATION NUMBER: 60/089598  
70 PRIOR FILING DATE: 1998-06-17  
71 PRIOR APPLICATION NUMBER: 60/089599  
72 PRIOR FILING DATE: 1998-06-17  
73 PRIOR APPLICATION NUMBER: 60/089600  
74 PRIOR FILING DATE: 1998-06-17  
75 PRIOR APPLICATION NUMBER: 60/089653  
76 PRIOR FILING DATE: 1998-06-17  
77 PRIOR APPLICATION NUMBER: 60/089801  
78 PRIOR FILING DATE: 1998-06-18  
79 PRIOR APPLICATION NUMBER: 60/089907  
80 PRIOR FILING DATE: 1998-06-18  
81 PRIOR APPLICATION NUMBER: 60/089908  
82 PRIOR FILING DATE: 1998-06-18  
83 PRIOR APPLICATION NUMBER: 60/089947  
84 PRIOR FILING DATE: 1998-06-19  
85 PRIOR APPLICATION NUMBER: 60/089948  
86 PRIOR FILING DATE: 1998-06-19  
87 PRIOR APPLICATION NUMBER: 60/089952  
88 PRIOR FILING DATE: 1998-06-19  
89 PRIOR APPLICATION NUMBER: 60/090246  
90 PRIOR FILING DATE: 1998-06-22  
91 PRIOR APPLICATION NUMBER: 60/090252  
92 PRIOR FILING DATE: 1998-06-22  
93 PRIOR APPLICATION NUMBER: 60/090254  
94 PRIOR FILING DATE: 1998-06-22  
95 PRIOR APPLICATION NUMBER: 60/090349  
96 PRIOR FILING DATE: 1998-06-23  
97 PRIOR APPLICATION NUMBER: 60/090355  
98 PRIOR FILING DATE: 1998-06-23  
99 PRIOR APPLICATION NUMBER: 60/090429  
100 PRIOR FILING DATE: 1998-06-24  
101 PRIOR APPLICATION NUMBER: 60/090431  
102 PRIOR FILING DATE: 1998-06-24  
103 PRIOR APPLICATION NUMBER: 60/090435  
104 PRIOR FILING DATE: 1998-06-24  
105 PRIOR APPLICATION NUMBER: 60/090444  
106 PRIOR FILING DATE: 1998-06-24  
107 PRIOR APPLICATION NUMBER: 60/090445  
108 PRIOR FILING DATE: 1998-06-24  
109 PRIOR APPLICATION NUMBER: 60/090472  
110 PRIOR FILING DATE: 1998-06-24  
111 PRIOR APPLICATION NUMBER: 60/090535  
112 PRIOR FILING DATE: 1998-06-24  
113 PRIOR APPLICATION NUMBER: 60/090576  
114 PRIOR FILING DATE: 1998-06-25  
115 PRIOR APPLICATION NUMBER: 60/090678  
116 PRIOR FILING DATE: 1998-06-25  
117 PRIOR APPLICATION NUMBER: 60/090690  
118 PRIOR FILING DATE: 1998-06-25  
119 PRIOR APPLICATION NUMBER: 60/090694  
120 PRIOR FILING DATE: 1998-06-25  
121 PRIOR APPLICATION NUMBER: 60/090695  
122 PRIOR FILING DATE: 1998-06-25  
123 PRIOR APPLICATION NUMBER: 60/090696  
124 PRIOR FILING DATE: 1998-06-25  
125 PRIOR APPLICATION NUMBER: 60/090862  
126 PRIOR FILING DATE: 1998-06-26  
127 PRIOR APPLICATION NUMBER: 60/090863  
128 PRIOR FILING DATE: 1998-06-26  
129 PRIOR APPLICATION NUMBER: 60/091360  
130 PRIOR FILING DATE: 1998-07-01  
131 PRIOR APPLICATION NUMBER: 60/091478  
132 PRIOR FILING DATE: 1998-07-02  
133 PRIOR APPLICATION NUMBER: 60/091544  
134 PRIOR FILING DATE: 1998-07-01  
135 PRIOR APPLICATION NUMBER: 60/091519  
136 PRIOR FILING DATE: 1998-07-02  
137 PRIOR APPLICATION NUMBER: 60/091626  
138 PRIOR FILING DATE: 1998-07-02  
139 PRIOR APPLICATION NUMBER: 60/091633  
140 PRIOR FILING DATE: 1998-07-02

;; PRIOR APPLICATION NUMBER: 60/091978  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/092182  
;; PRIOR FILING DATE: 1998-07-09

```
Query Match      99.3%; Score 4719.5; DB 4; Length 867;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 866; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1 MGPPSLVLCCLLSATVFLSGSSAFSLSHRLKGRFQDRNRIRNIILVLTDDQVGLGS 60
DB 1 MGPPSLVLCCLLSATVFLSGSSAFSLSHRLKGRFQDRNRIRNIILVLTDDQVGLGS 60

QY 61 MQVMNKTIRIMEQGGTHFINAFVTPPMCCPSRSSILTKGYVHNHTYNNENSSPSWQA 120
DB 61 MQVMNKTIRIMEQGGTHFINAFVTPPMCCPSRSSILTKGYVHNHTYNNENSSPSWQA 120

QY 121 QHESRTFAVLNMGTCYRTAFCKYLYNEVSGSVPPGKWKVGLLKNRSFYNTLCRNGVK 180
DB 121 QHESRTFAVLNMGTCYRTAFCKYLYNEVSGSVPPGKWKVGLLKNRSFYNTLCRNGVK 180

QY 181 EXHGSYSDKDYLTDLITNDSVSFFRTSKOMYPHRPVLVISHAAPHGPDSPAPQYSRLFP 240
DB 181 EXHGSYSDKDYLTDLITNDSVSFFRTSKOMYPHRPVLVISHAAPHGPDSPAPQYSRLFP 240

QY 241 NASOHTITSYVAPNPDKHMIRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
DB 241 NASOHTITSYVAPNPDKHMIRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300

QY 301 LVETGELONTYIVYTADHGHIHQFGLVKGSMPEYFDIRVPFVVRGPNVEAGCLNPHIV 360
DB 301 LVETGELONTYIVYTADHGHIHQFGLVKGSMPEYFDIRVPFVVRGPNVEAGCLNPHIV 360

QY 361 LNIDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRFHLKKQVWRDSFLVERGKLL 420
DB 361 LNIDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRFHLKKQVWRDSFLVERGKLL 420

QY 421 HKRDNKVDQAEEENFLPKYQVVKOLCORAEYQTAQCEQLGQKMQCQVEDATGKLLHKCKGP 480
DB 421 HKRDNKVDQAEEENFLPKYQVVKOLCORAEYQTAQCEQLGQKMQCQVEDATGKLLHKCKGP 480

QY 481 MRLGGSRALSNLVPKYQGGSEACTCDSGDYKLSLAGRRKKLPKKYKASYVRSRSIRSV 540
DB 481 MRLGGSRALSNLVPKYQGGSEACTCDSGDYKLSLAGRRKKLPKKYKASYVRSRSIRSV 540

QY 541 ALEVDRVYTHVGLGDAAPRNLTKEHWPAGPEDDQDDKGGDFSGTGGLPDYSAANPIKVT 600
DB 541 ALEVDRVYTHVGLGDAAPRNLTKEHWPAGPEDDQDDKGGDFSGTGGLPDYSAANPIKVT 600

QY 601 HRCYILENDTVQCDLDLYKSLQAWKDKHLHIDHEIETLQNKIKNLEVRGHLKKRPEEC 660
DB 601 HRCYILENDTVQCDLDLYKSLQAWKDKHLHIDHEIETLQNKIKNLEVRGHLKKRPEEC 660

QY 661 DCHKISYHTQHKGRLLKRGSSLHPFRKGLQEKDKVWLLRQKRRKKLLKRLQNNNDTC 720
DB 661 DCHKISYHTQHKGRLLKRGSSLHPFRKGLQEKDKVWLLRQKRRKKLLKRLQNNNDTC 720

QY 721 SMPGLTCTFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNFCEPATGFLE 780
DB 721 SMPGLTCTFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNFCEPATGFLE 780

QY 781 YFDLNTDPVQLMNAVNTLDRVNLQHLVQJLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840
DB 781 YFDLNTDPVQLMNAVNTLDRVNLQHLVQJLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840

QY 841 QYRQFORRWKPEWKPSSKSLGQLEWEGEG 870
DB 838 QYRQFORRWKPEWKPSSKSLGQLEWEGEG 867
```

US-09-886-319A-36  
; Sequence 36, Application US/09886319A  
; Patent No. 6586185  
; GENERAL INFORMATION:  
; APPLICANT: Wolf, Eckard  
; APPLICANT: Werner, Sabine  
; APPLICANT: Halle, Jörn-Peter  
; APPLICANT: Regenbogen, Johannes  
; APPLICANT: Coppel, Andreas  
; TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for  
; TITLE OF INVENTION: the Diagnosis or Treatment of Skin Disorders and Wound  
; TITLE OF INVENTION: Healing and for the Identification of Pharmacologically  
; TITLE OF INVENTION: Active Substances  
; FILE REFERENCE: 50125/014002  
; CURRENT APPLICATION NUMBER: US/09/886,319A  
; CURRENT FILING DATE: 2001-06-20  
; PRIOR APPLICATION NUMBER: US 60/222,081  
; PRIOR FILING DATE: 2000-08-01  
; PRIOR APPLICATION NUMBER: DE 10030149.5  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 84  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 36  
; LENGTH: 690  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-886-319A-36

```
Query Match      79.5%; Score 3776; DB 4; Length 690;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 181 EXHGSYSDKDYLTDLITNDSVSFFRTSKOMYPHRPVLVISHAAPHGPDSPAPQYSRLFP 240
DB 1 EXHGSYSDKDYLTDLITNDSVSFFRTSKOMYPHRPVLVISHAAPHGPDSPAPQYSRLFP 60

QY 241 NASOHTITSYVAPNPDKHMIRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
DB 61 NASOHTITSYVAPNPDKHMIRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 120

QY 301 LVETGELONTYIVYTADHGHIHQFGLVKGSMPEYFDIRVPFVVRGPNVEAGCLNPHIV 360
DB 121 LVETGELONTYIVYTADHGHIHQFGLVKGSMPEYFDIRVPFVVRGPNVEAGCLNPHIV 180

QY 361 LNIDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRFHLKKQVWRDSFLVERGKLL 420
DB 181 LNIDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRFHLKKQVWRDSFLVERGKLL 240

QY 421 HKRDNKVDQAEEENFLPKYQVVKOLCORAEYQTAQCEQLGQKMQCQVEDATGKLLHKCKGP 480
DB 241 HKRDNKVDQAEEENFLPKYQVVKOLCORAEYQTAQCEQLGQKMQCQVEDATGKLLHKCKGP 300

QY 481 MRLGGSRALSNLVPKYQGGSEACTCDSGDYKLSLAGRRKKLPKKYKASYVRSRSIRSV 540
DB 301 MRLGGSRALSNLVPKYQGGSEACTCDSGDYKLSLAGRRKKLPKKYKASYVRSRSIRSV 360

QY 541 ALEVDRVYTHVGLGDAAPRNLTKEHWPAGPEDDQDDKGGDFSGTGGLPDYSAANPIKVT 600
DB 361 ALEVDRVYTHVGLGDAAPRNLTKEHWPAGPEDDQDDKGGDFSGTGGLPDYSAANPIKVT 420

QY 601 HRCYILENDTVQCDLDLYKSLQAWKDKHLHIDHEIETLQNKIKNLEVRGHLKKRPEEC 660
DB 421 HRCYILENDTVQCDLDLYKSLQAWKDKHLHIDHEIETLQNKIKNLEVRGHLKKRPEEC 480

QY 661 DCHKISYHTQHKGRLLKRGSSLHPFRKGLQEKDKVWLLRQKRRKKLLKRLQNNNDTC 720
DB 481 DCHKISYHTQHKGRLLKRGSSLHPFRKGLQEKDKVWLLRQKRRKKLLKRLQNNNDTC 540

QY 721 SMPGLTCTFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNFCEPATGFLE 780
DB 541 SMPGLTCTFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNFCEPATGFLE 600

QY 781 YFDLNTDPVQLMNAVNTLDRVNLQHLVQJLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840
```

Db 601 YFDLNDTPYOLMNAVNTLDRDVLNQLHVQLMELRSCKGYKCNPRTRNMDLGLKGGSYE 660  
Qy 841 QYRQORRKPMPKRPSSKSLQLMWEGWG 870  
Db 661 QYRQORRKPMPKRPSSKSLQLMWEGWG 690

RESULT 3  
US-09-773-426A-3  
; Sequence 3, Application US/09773426A  
; Patent No. 6534302  
; GENERAL INFORMATION:  
; APPLICANT: Gluckman, Maria Alexandra  
; APPLICANT: Williamson, Mark  
; APPLICANT: Teia, Fong-Ying  
; APPLICANT: Rudolph-Owen, Laura A.  
; TITLE OF INVENTION: 22438, 23553, 25278, and 26212 No. 6534302e1  
; FILE OF INVENTION: Human Sulfatases (A CIP Application)  
; CURRENT FILING DATE: 2001-01-31  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: US 09/495,823  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 871  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-773-426A-3

Query Match 66.8%; Score 3175.5; DB 4; Length 871;  
Best Local Similarity 64.9%; Pred. No. 3.1e-301;  
Matches 578; Conservative 120; Mismatches 142; Indels 51; Gaps 12;

Qy 5 SLVLCILSATVFSLLGGSSAFSLSHRLKGRFORDRNRIRPNILVLTDDQDVELGSMQVM 64  
Db 7 ALVLAVLGT---ELLSGLSTVRSFRGRIOQRKRNIRPNILVLTDDQDVELGSLQVM 63  
Qy 65 NKTRIMEOGGTHFINAFVTTMCCPSRSSILTKYVHNHNTYTNENCSPPSQAQHE 124  
Db 64 NKTRIMEGGATFINAFVTTMCCPSRSSMLTKYVHNHNTYTNENCSPPSQAQHEP 123  
Qy 125 RFAVYLNSTGYRTAFPGKYLNEYSGVYPPGKEMVGLLKNRSFYNTLCRNGVKEKHG 184  
Db 124 RFAVYLNSTGYRTAFPGKYLNEYSGVYPPGKEMVGLLKNRSFYNTLCRNGVKEKHG 183  
Qy 185 SLYSKDYLTLTNDVSVFRTSKMYPHRLVWVISHAAPHGSDSAPQYSLRFPNASQ 244  
Db 184 FOYAKDYFTDLITNESINVFMSKGIYHPHPIWVISHAAPHGSDSAPQYSLRFPNASQ 243  
Qy 245 HITPSYNYAPNDKHWIMQYTGPMPIHMEFTNMLORKELOTLMSVDDSMETIYNMLVET 304  
Db 244 HITPSYNYAPNDKHWIMQYTGPMPIHMEFTNMLORKELOTLMSVDDSMETIYNMLVET 303  
Qy 305 GELDNTYIVYTADHGVIHQFGLVKGKSPYBFDIRVPFYVRGPNVAGCLNPHIVLNI 364  
Db 304 GELENTYIIVYTADHGVIHQFGLVKGKSPYBFDIRVPFYVRGPNVAGCLNPHIVLNI 363  
Qy 365 LAPTIIDLAGLDIPADMOKSILKLDTRPVRNPHILKQKRWVDSFIVERGKLLHQRD 424  
Db 364 LAPTIIDLAGLDIPADMOKSILKLDTRPVRNPHILKQKRWVDSFIVERGKLLHQRD 423  
Qy 425 NDKVDAQEEFLPKYORVNDLCORAEVQATCEGLQKQOCVEDATGKGLKHLKCKGPMELG 484  
Db 424 ESKNIOOSNHLPKYERVELCOQARYQATCOPGQKQOCIEDTSGLKRIHCKGSPDLL 483  
Qy 485 GSR-ALSNLVPKYGGSEACTCDSDGYKLSLAGRR-KKLFPK-----KYKASYVRSR 537  
Db 484 TVRQSTNLYARGFHDKDKESCKESGYSRASRSQRKSLRQQLRNQOTPKYKPRFVHTROT 543  
Qy 538 RSVAIEDGVRHVHVLGD-----AAQPNLITKEHNEG--APEDQDDKGG-----DFS 583

Db 544 RSLSEFEGEIYDINLEBEEBELQVLQPRNTAKRHDEGHKGRDQLQASSGGNRRGRLADSS 603  
Qy 584 GTGGGLPDYSAANPIKVTYHRCYILENDTVOCDDLYKSLQAKWKHKLHDIETLQNKIK 643  
Db 604 NAVGPP-----TTVRVTHKCFILPNDISHCBRELYQSARAWKDKAYIDKEIEALQDKIK 658  
Qy 644 NLREVRGHLKCKRPEECCHKISYHTQHKGRKLR--GSSLHPRFKGLQKDK-KVMILLRE 700  
Db 659 NLREVRGHLKCKRPEECSCSKSYNKEGKVKQEKLSHLHPEKEAAQEVDSKLQFKKE 718  
Qy 701 -OKRKKKRLKRLKRLQNDTCSMPGLCTCTHDNHOHQTAPFTWLTGPFCACTSANNTYWC 759  
Db 719 NNRRKKRKKRKKRQKRGKSGSLFGLCTCTHDNHOHQTAPFTWLTGPFCACTSANNTYWC 778  
Qy 760 MRTINETHFLFCEPATGFLFYFDLNTDPYOLMNAVNTLDRDVLNQLHVQLMELRSCKGY 819  
Db 779 LRTVNETHFLFCEPATGFLFYFDLNTDPYOLMNAVNTLDRDVLNQLHVQLMELRSCKGY 838  
Qy 820 KOCNPRTRNMDLGLKGGSYEQYQFQRRKWPMPKRPSSKSLQLMWEGWG 870  
Db 839 KOCNPRPKNLDVGNKGGSYDLHR-----GQLMDGWEG 871

RESULT 4  
US-09-668-673B-2  
; Sequence 2, Application US/09668673B  
; Patent No. 6562956  
; GENERAL INFORMATION:  
; APPLICANT: Emerson, Charles P  
; APPLICANT: Dhoot, Gurtej K  
; TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF  
; TITLE OF INVENTION: SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR  
; TITLE OF INVENTION: CHARACTERIZATION OF SUCH PROTEINS  
; FILE REFERENCE: PENN-0733  
; CURRENT APPLICATION NUMBER: US/09/668.673B  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155,738  
; PRIOR FILING DATE: 1999-09-23  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 867  
; TYPE: PRT  
; ORGANISM: quails  
US-09-668-673B-2

Query Match 65.6%; Score 3117.5; DB 4; Length 867;  
Best Local Similarity 63.9%; Pred. No. 1.5e-295;  
Matches 571; Conservative 117; Mismatches 146; Indels 59; Gaps 10;

Qy 5 SLVLCILSATVFSLLGGSSAFSLSHRLKGRFORDRNRIRPNILVLTDDQDVELGSMQVM 64  
Db 7 ALVLAVLST---ELTSSHSSTLKSILRFRGRVQQRKRNIRPNILVLTDDQDVELGSLQVM 63  
Qy 65 NKTRIMEOGGTHFINAFVTTMCCPSRSSILTKYVHNHNTYTNENCSPPSQAQHE 124  
Db 64 NKTRIMEGGASFINAFVTTMCCPSRSSMLTKYVHNHNTYTNENCSPPSQAQHEP 123  
Qy 125 RFAVYLNSTGYRTAFPGKYLNEYSGVYPPGKEMVGLLKNRSFYNTLCRNGVKEKHG 184  
Db 124 RFAVYLNSTGYRTAFPGKYLNEYSGVYPPGKEMVGLLKNRSFYNTLCRNGVKEKHG 183  
Qy 185 SLYSKDYLTLTNDVSVFRTSKMYPHRLVWVISHAAPHGSDSAPQYSLRFPNASQ 244  
Db 184 FOYAKDYFTDLITNESINVFMSKGIYHPHPIWVISHAAPHGSDSAPQYSLRFPNASQ 243  
Qy 245 HITPSYNYAPNDKHWIMQYTGPMPIHMEFTNMLORKELOTLMSVDDSMETIYNMLVET 304  
Db 244 HITPSYNYAPNDKHWIMQYTGPMPIHMEFTNMLORKELOTLMSVDDSMETIYNMLVET 303  
Qy 305 GELDNTYIVYTADHGVIHQFGLVKGKSPYBFDIRVPFYVRGPNVAGCLNPHIVLNI 364

```
Db 304 GELENTYIIYTADHGVIHQFGLVKGKSMYDFDIRVPPFFIRGPSVPGSVVPQIVLNID 363
QY 365 LAPTILDIAGLIDIPADMGKSLKLLDTERPVNRPHLKKQKRVWRDPSFLVERGKLKHRD 424
Db 364 LAPTILDIAGLIDTPADMGKSLKLLDTERPGRNFRNTNKTIKWRDTFLVERGKFLRKE 423
QY 425 NDKVDAQEBEFLPKYQKRVKDLQCBAYOTACBOLGQKQCVEDATGKLKHLKHKGPMRLG 484
Db 424 EANKWTQSQNLQPKYKRVKQELQCBAYOTACBOLGQKQCVEDATGKLKHLKHKGPMRLG 483
QY 485 GSRALSNLY-PKYGGGSEACTCDSDGKLSLAGRR-----KKLFFKKYKASYVRSRSI 537
Db 484 ATRKRTSRHSRGYSGKDKCNCGDTDFNSRTQRKNQQLFLNPSAQKYYKPRFVHTRTQ 543
QY 538 RSVAEVDRGVYHVLGD---AAQPRNLTKKH-----WPGAPEDODDQGGGDFS 583
Db 544 RLSVSEFEIEIYDINLEEELOVLKTRSTIKKHNAENDKKAETDGPAGDTMVGDTDI 603
QY 584 GTGGLPDYSAANPIKYTHRCYILENDTVQCDLDLYKSLQAKMDKHLHIDHEIETLQNKIK 643
Db 604 G-----QPSSRVYTHKCFILPNDTIRCEBELYQSARAKMDHKAYIDKIEALQDKIK 655
QY 644 NLREVRGHLKRRPREDCCHKISYHTQHG-----RLKHRSSSLHPPFKGLQEKD-KVWL 697
Db 656 NLREVRGHLKRRPREDCCHKISYHTQHG-----SHLHPFKBAQAQEVDSKLQL 712
QY 698 LREOKKKKLLKLLKRLQNNDTCSMPLGTCFTHDQHMOTAPFWTLGPFCACTSANNTY 757
Db 713 FXENRRKKERKGGKQKGGSCSLGTLCTFDHNNHWTAPFWTLGPFCACTSSNNNTY 772
QY 758 WCMRTINETHFLCFEATGFLFYDLNTDTPYQLMNAVNTLDRVNLQHLVQLMELRSCK 817
Db 773 WCLRTVNDTHNLFCEPATGFLFEFDMNTDTPYQLMNAVNTLDRVNLQHLVQLMELRSCK 832
QY 818 GYKQCNPRFNDLGLKQGSYEQVRQRRKQKPEKPSKSLGOLWEGEG 870
Db 833 GYKQCNPRFNDLGLKQGSYEQVRQRRKQKPEKPSKSLGOLWEGEG 867

RESULT 5
US-09-668-673B-16
; Sequence 16, Application US/09668673B
; Patent No. 6562956
; GENERAL INFORMATION:
; APPLICANT: Emerson, Charles P
; TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF
; SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR
; CHARACTERIZATION OF SUCH PROTEINS
; FILE REFERENCE: PERN-0733
; CURRENT APPLICATION NUMBER: US/09/668,673B
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,738
; PRIOR FILING DATE: 1999-09-23
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 1611
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (819)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (840)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (844)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (852)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (858)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (865)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (875)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (881)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (888)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (896)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (907)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (910)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (915)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (927)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (943)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (945)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (948)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (954)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (959)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (971)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
```

```

; LOCATION: (974)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1018)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1046)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1080)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1089)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1102)..(1103)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1105)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1121)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1127)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1191)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1199)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1223)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1235)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1250)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1307)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1321)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1356)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1362)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1382)..(1383)

```

```

; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1397)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1431)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1437)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1448)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1458)
; OTHER INFORMATION: any amino acid
; FEATURE:

Query Match      64.4%; Score 3059; DB 4; Length 1611;
Best Local Similarity 65.6%; Pred. No. 2.1e-289;
Matches 552; Conservative 112; Mismatched 129; Indels 48; Gaps 11;

QY 55 DVELGSMQVANKTRIMEQGTHFINAFVTTMCCPSRSSILTKGYVHNHNTYNNENCS 114
Db 1 DVELGSLQVMNKRIMEHGGATFINAFVTTMCCPSRSSMLTKGYVHNHNTYNNENCS 60

QY 115 SPSWQAHESRTFAVYLNSTGYRTAFPGKYLNEYNGSYVPCWKEMVGLLKNREYNTL 174
Db 61 SPSWQAHHEPRTFAVYLNNTGYRTAFPGKYLNEYNGSYVPCWREWLGLIKNSRPYNTV 120

QY 175 CRNGVKEHGSYSDYKDLTDLITNDSVSFFRTSKQYRHPVLMVISHAAPHGSDSAPQ 234
Db 121 CRNGIKHGFYAKDYFTDLITNESINYFKMSKMYRHPVPMVISHAAPHGSDSAPQ 180

QY 235 YSLFPNASCHITPSYNYAPNPKWIMRYTQPMKPIHMEFTNMLQRLQTLMSVDSM 294
Db 181 FSKLYPNASCHITPSYNYAPNMDKHWIMQYTGPMPLPIHMEFTNMLQRLQTLMSVDSV 240

QY 295 ETIYNMLVETGELDNTYIYTTADHGVHIGQFGLVKGSMYPYEFDIRVPYVRGPNVEAGC 354
Db 241 ERLYNMLVETGELENTYIYTTADHGVHIGQFGLVKGSMYPYDFDIRVPFFTEGSPVEGS 300

QY 355 LNPHIVINIDLAPTILDIAGLDI PADMDGKSLKLLDTERPVRNPHLKKQMRVWDSFLV 414
Db 301 IVPQIVLINIDLAPTILDIAGLDTPPDVDGKSLKLLDPEKPGNRFRNTKCAKIMRDTFLV 360

QY 415 ERGKLLHKRDNDKVDQAEENFLPKYQVKDLCORAEYOTACEQLGQKWCQVEDATGKLKL 474
Db 361 ERGKFLRKEESSKNIQSNHLPKRYVELQCOQARYQTACEQPCQKWCQIEDTSGKLRI 420

QY 475 HKCKGPMRLGSSR-ALSNLVPKYIYQGSSEACTCDSDGYKLSLAGRR-KKLFK-----KY 527
Db 421 HKCKGPSDLLTVRQSTNLYARGFHDKDKCESGYSRASRSQKSKORQLRQNGTPKY 480

QY 528 KASYVRSRSIRSVAIENDVRVTHVGLD-----AAQPNLTGHWPG--APEDQDDKQGG 580
Db 481 KPRFVHTQTRSLSEFEIGEYIDNLEBEEELQVLPQPNIAKRHDEGHGKPRDLQASSGG 540

QY 581 -----DFSGTGGLPDYSAANPIKVTHRCYILENDTVQCOLDYKSLQAKHKLHLDH 633
Db 541 NRCGMLADSSNAVGPP-----TTVRVTHKCFILPNDSTHCEBELYQSARANKHAYIDK 595

QY 634 ELETQNKIKNLRVRGHLKXKRPBECCHKISYHTQHKGLKGR--GSSLHPRFKGQJB 691
Db 596 EIEALQDKIKNLRVRGHLKXKRPBECSCSKQSYNKEKGKVKQEKKSLHLPFKAAQE 655

QY 692 KD-KWLLRE-QKRBKKGLKLLKRLQNNDTCSMPLGTCFTHDNQHWOTAPFWTLGPFAC 749

```

Db 656 VDSKQLFLKNNRRKKERKEKRQRKGBCSLPGLTCTHNNHQTAPFWNLGSFCA 715  
 QY 750 TSANNNTYMCWRTTINETHNLFCEPATGLEYEDLNTDYPQLMNAVNTLDRDVLNQLHVO 809  
 Db 716 TSSNNNTYMCWRTTINETHNLFCEPATGLEYEDLNTDYPQLMNAVNTLDRDVLNQLHVO 775  
 QY 810 LMELRSCGKYKQCNPRTRNMDLGLKDGGSYEQYQFQRKWPMPKPSKSLGQLWEGWE 869  
 Db 776 LMELRSCGKYKQCNPRTRNMDLGLKDGGSYEQYQFQRKWPMPKPSKSLGQLWEGWE 817  
 QY 870 G 870  
 Db 818 G 818

RESULT 6  
 US-09-668-673B-14  
 ; Sequence 14, Application US/09668673B  
 ; Patent No. 6562956  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Emerson, Charles P  
 ; APPLICANT: Dhoot, Gurtej K  
 ; TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF  
 ; TITLE OF INVENTION: SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR  
 ; TITLE OF INVENTION: CHARACTERIZATION OF SUCH PROTEINS  
 ; FILE REFERENCE: PENN-0733  
 ; CURRENT APPLICATION NUMBER: US/09/668,673B  
 ; CURRENT FILING DATE: 2000-09-22  
 ; PRIOR APPLICATION NUMBER: 60/155,738  
 ; PRIOR FILING DATE: 1999-09-23  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 14  
 ; LENGTH: 455  
 ; TYPE: PRT  
 ; ORGANISM: Mus sp.  
 ; FEATURE:  
 ; NAME/KEY: UNSURE  
 ; LOCATION: (142-143)  
 ; OTHER INFORMATION: any amino acid  
 ; FEATURE:  
 ; NAME/KEY: UNSURE  
 ; LOCATION: (146)  
 ; OTHER INFORMATION: any amino acid  
 ; FEATURE:  
 ; NAME/KEY: UNSURE  
 ; LOCATION: (445)  
 ; OTHER INFORMATION: any amino acid  
 ; OTHER INFORMATION: any amino acid  
 US-09-668-673B-14

Query Match 44.1%; Score 2093; DB 4; Length 455;  
 Best Local Similarity 86.1%; Pred. No. 8.3e-196;  
 Matches 384; Conservative 16; Mismatches 38; Indels 8; Gaps 4;  
 QY 431 QEEFLPKYQVVKDCORAEYQTACEQGGKQWQVCEATGKLLKCKGKPMEL---GGSR 487  
 Db 1 QEEFLPKYQVVKDCORAEYQTACEQGGKQWQVCEATGKLLKCKGKPMELFGGGGSR 60  
 QY 488 ALSNLVPKYQGSSEACTCDSD---GDYKLSLAGRRKKLPKQYKASYVRSRSIRSAIEV 544  
 Db 61 ALSNLVPKYQGSSEACTCDSDGGGDDYKLAGRR-KLFKKYKTSYARNRSIRSAIEV 119  
 QY 545 DGRVYHVLGDAAPRNLTKEHWPAPEDODDQGGDFSGTGLPDYSAANPIKVTHRCY 604  
 Db 120 DGEIYHVLGDTVPPRNLSKPHXGAXEDQDDKDGSGFSGTGLPDYSAANPIKVTHRCY 179  
 QY 605 ILENDTVQCDLGLYKSLQAWKHKLHIDHEIETLQNKIKNLEVRGHLKQKPEECDCCHK 664  
 Db 180 ILENDTVQCDLGLYKSLQAWKHKLHIDHEIETLQNKIKNLEVRGHLKQKPEECDCCHK 239  
 QY 665 ISYTHQKGRUKHKGSSILHPRKGLQEKDKVLLRQKQKQKURKLLKRLQNNDCSMPG 724  
 Db 240 ISYSHQKGRUKHKGSSILHPRKGLQEKDKVLLRQKQKQKURKLLRQNNDCSMPG 298

QY 725 LTCFTHDNOHQTAPFWTLGPFCACSTANNNTYMCWRTTINETHNLFCEPATGLEYEDL 784  
 Db 299 LTCFTHDNOHQTAPFWTLGPFCACSTANNNTYMCWRTTINETHNLFCEPATGLEYEDL 358  
 QY 785 NTDPQLMNAVNTLDRDVLNQLHVMELRSCKYKQCNPRTRNMDLGLKDGGSYEQYRQ 844  
 Db 359 STDPQLMNAVNTLDRDVLNQLHVMELRSCKYKQCNPRTRNMDLGLKDGGSYEQYRQ 418  
 QY 845 FORRKWPEMKRPPSSKSLGQLWEGWEG 870  
 Db 419 FORRKWPEMKRPPSSKSLGQLWEGWEG 444

RESULT 7  
 US-09-886-319A-35  
 ; Sequence 35, Application US/09886319A  
 ; Patent No. 6586185  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wolf, Eckard  
 ; APPLICANT: Werner, Sabine  
 ; APPLICANT: Halle, Jorn-Peter  
 ; APPLICANT: Regenbogen, Johannes  
 ; APPLICANT: Goppelt, Andreas  
 ; TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for  
 ; TITLE OF INVENTION: the Diagnosis or Treatment of Skin Disorders and Wound  
 ; TITLE OF INVENTION: Healing and for the Identification of Pharmacologically  
 ; TITLE OF INVENTION: Active Substances  
 ; FILE REFERENCE: 50125/014002  
 ; CURRENT APPLICATION NUMBER: US/09/886,319A  
 ; CURRENT FILING DATE: 2001-06-20  
 ; PRIOR APPLICATION NUMBER: US 60/222,081  
 ; PRIOR FILING DATE: 2000-08-01  
 ; PRIOR APPLICATION NUMBER: DE 10030149.5  
 ; PRIOR FILING DATE: 2000-06-20  
 ; NUMBER OF SEQ ID NOS: 84  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 35  
 ; LENGTH: 470  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 ; FEATURE:  
 ; NAME/KEY: VARIANT  
 ; LOCATION: 371  
 ; OTHER INFORMATION: Xaa= any amino acid  
 US-09-886-319A-35

Query Match 35.8%; Score 1702; DB 4; Length 470;  
 Best Local Similarity 75.3%; Pred. No. 1.5e-157;  
 Matches 327; Conservative 21; Mismatches 36; Indels 50; Gaps 7;  
 QY 300 MLVETGELDNTYIVYTADGHYHIGQGLVKGKSMYPEDIRVPFYVRGPNVYAGCLNPHI 359  
 Db 1 MLVETGELDNTYIVYTADGHYHIGQGLVKGKSMYPEDIRVPFYVRGPNVYAGCLNPHI 60  
 QY 360 VLNIDLAP-----TILDIAGLDIPADMDGKSLKLLDTEPVN 397  
 Db 61 VLNIDLAPHTGYRTWGPCRHGREVYSQTTGLRAASEPVPLEKEAEGL-----GETP-- 113  
 QY 398 RFELKQKQVWRDSDFLVERGKLLKHKRDNDKYDAQBEENFLPKYQVVKDCORAEYQTACEQ 457  
 Db 114 -----SWMREA-----NLLHKREGDKVNAQBEENFLPKYQVVKDCORAEYQTACEQ 159  
 QY 458 LQKQWQCVEDATGKLLKHKCKGPMEL-----GGSRALSNLVPKYQGSSEACTCDSD---GDY 511  
 Db 160 LQKQWQCVEDATGKLLKHKCKGPMELFGGGGSRALSNLVPKYQGSSEACTCDSDGGGGDY 219  
 QY 512 KLSLAGRRKKLPKQYKASYVRSRSIRSAIEVQGRVYHVLGDAAPRNLTKEHWPAP 571  
 Db 220 KLSLAGRR-KLPKQYKTSYARNRSIRSAIEVQGEIYHVLGDTVPPRNLSKQHWEGAP 278  
 QY 572 EQDDKDGDFSGTGLPDYSAANPIKVTHRCYILENDTVQCDLGLYKSLQAWKHKLHI 631

Db 279 EDQDDKGGFSGTGLPLDYSPNPIKVTHRCVILENDTVQCOLDLYKSLQAKDKHLHI 338  
Qy 632 DHEIETLQNKILNREVRGHLKKRPECDCCHKISVHTQHGRLKRGCSLHPFRKGLQE 691  
Db 339 DHEIETLQNKILNREVRGHLKKRPECDCCHKISVHSQHKGRXKHKGSSLHPFRKGLQE 398  
Qy 692 KDKVLLREKQKRXK 705  
Db 399 KDKVLLRDRNRR 412  
RESULT 8  
US-09-810-347-4  
; Sequence 4, Application US/09810347  
; Patent No. 6461847  
; GENERAL INFORMATION:  
; APPLICANT: YE, Jane et al.  
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001169  
; CURRENT APPLICATION NUMBER: US/09/810,347  
; CURRENT FILING DATE: 2001-03-19  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 360  
; TYPE: PRT  
; ORGANISM: Human  
US-09-810-347-4

Query Match 33.8%; Score 1606.5; DB 4; Length 360;  
Best Local Similarity 79.8%; Pred. No. 2.2e-148; Indels 3; Gaps 1;  
Matches 285; Conservative 40; Mismatches 29;  
Qy 5 SILVCLLSATVFSLLGGSSAFLSHRLKGRFQRRNIRPNILVLTDDQDVELGSMQVM 64  
Db 7 ALVLAVLGT---ELLGSLCSTVRSFPRGRIGQERKNIRPNILVLTDDQDVELGSLQVM 63  
Qy 65 NKTRIMEQGGTHFINAFVTTMCCPSSSILTKYVHNHTYNNENSCSSPSQAOAHES 124  
Db 64 NKTRIMEHGATFINAFVTTMCCPSSSMLTKYVHNHTYNNENSCSSPSQAOAHPE 123  
Qy 125 RTFAVILNSTGYRTAFPGKYLNEYNGSVYPPGKWEVGLLKNRFRNYTLCRNGVKEKHG 184  
Db 124 RTFAVILNSTGYRTAFPGKYLNEYNGSVYPPGKWEVGLLKNRFRNYTLCRNGVKEKHG 183  
Qy 185 SDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAPHGPDSSAPQYSLFPNASQ 244  
Db 184 FDYAKDYFTDLITNESINYPFKSKMYPHRPVLMVISHAAPHGPDSSAPQYSLFPNASQ 243  
Qy 245 HITPSYNYAPNPKDKNIMRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNMLVET 304  
Db 244 HITPSYNYAPNPKDKNIMRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNMLVET 303  
Qy 305 GELDNTYIVYTDHGYHIGQGLVKGKSNPYEFDIRVPFYVGRGPNVBAAGCL 361  
Db 304 GLENTYIITADHGYHIGQGLVKGKSNPYEFDIRVPFFIRGSPVPGSIVPQIVL 360

RESULT 9  
US-09-810-347-2  
; Sequence 2, Application US/09810347  
; Patent No. 6461847  
; GENERAL INFORMATION:  
; APPLICANT: YE, Jane et al.  
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001169  
; CURRENT APPLICATION NUMBER: US/09/810,347  
; CURRENT FILING DATE: 2001-03-19  
; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 372  
; TYPE: PRT  
; ORGANISM: Human  
US-09-810-347-2

Query Match 33.5%; Score 1590.5; DB 4; Length 372;  
Best Local Similarity 80.1%; Pred. No. 8.4e-147; Indels 3; Gaps 1;  
Matches 281; Conservative 40; Mismatches 27;  
Qy 5 SILVCLLSATVFSLLGGSSAFLSHRLKGRFQRRNIRPNILVLTDDQDVELGSMQVM 64  
Db 7 ALVLAVLGT---ELLGSLCSTVRSFPRGRIGQERKNIRPNILVLTDDQDVELGSLQVM 63  
Qy 65 NKTRIMEQGGTHFINAFVTTMCCPSSSILTKYVHNHTYNNENSCSSPSQAOAHES 124  
Db 64 NKTRIMEHGATFINAFVTTMCCPSSSMLTKYVHNHTYNNENSCSSPSQAOAHPE 123  
Qy 125 RTFAVILNSTGYRTAFPGKYLNEYNGSVYPPGKWEVGLLKNRFRNYTLCRNGVKEKHG 184  
Db 124 RTFAVILNSTGYRTAFPGKYLNEYNGSVYPPGKWEVGLLKNRFRNYTLCRNGVKEKHG 183  
Qy 185 SDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAPHGPDSSAPQYSLFPNASQ 244  
Db 184 FDYAKDYFTDLITNESINYPFKSKMYPHRPVLMVISHAAPHGPDSSAPQYSLFPNASQ 243  
Qy 245 HITPSYNYAPNPKDKNIMRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNMLVET 304  
Db 244 HITPSYNYAPNPKDKNIMRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNMLVET 303  
Qy 305 GELDNTYIVYTDHGYHIGQGLVKGKSNPYEFDIRVPFYVGRGPNVBAAGCL 355  
Db 304 GLENTYIITADHGYHIGQGLVKGKSNPYEFDIRVPFFIRGSPVPGSIVPQIVL 354

RESULT 10  
US-09-668-673B-3  
; Sequence 3, Application US/09668673B  
; Patent No. 6562956  
; GENERAL INFORMATION:  
; APPLICANT: Emerson, Charles P  
; APPLICANT: Droot, Gurtej K  
; TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF  
; TITLE OF INVENTION: SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR  
; TITLE OF INVENTION: CHARACTERIZATION OF SUCH PROTEINS  
; FILE REFERENCE: PENN-0733  
; CURRENT APPLICATION NUMBER: US/09/668,673B  
; CURRENT FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155,738  
; PRIOR FILING DATE: 1999-09-23  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 709  
; TYPE: PRT  
; ORGANISM: Caenorhabditis elegans  
US-09-668-673B-3

Query Match 33.4%; Score 1589; DB 4; Length 709;  
Best Local Similarity 40.2%; Pred. No. 3.4e-146; Indels 158; Gaps 17;  
Matches 317; Conservative 118; Mismatches 196;  
Qy 45 NILVLTDDQDVELGSMQVMNKTTRIMEQGGTHFINAFVTTMCCPSSSILTKYVHNH 104  
Db 36 NVILILTDDQDIELGSMDFMPKTSQIMKRGTEFTSGYVTTPICCPSRSTILTGLYVHNH 95  
Qy 105 NTYNNENSCSSPSQAOAHESRTFAVILNSTGYRTAFPGKYLNEYNGSVYPPGKWEVGLL 164  
Db 96 HVATNNQNCCTGVEWRKVHKKSGIVYQAGVYRTAYGLKYLNEYNGSVYPPGWDSEWHAIV 155  
Qy 165 KNSFRNYTLCRNGVKEKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAA 224





US-08-484-493-11  
; Sequence 11, Application US/08484493  
; Patent No. 5728381  
; GENERAL INFORMATION:  
; APPLICANT: Wilson, Peter J  
; APPLICANT: Morris, Charles P  
; APPLICANT: Anson, Donald S  
; APPLICANT: Occhiodoro, Teresa  
; APPLICANT: Bielsicki, Julie  
; APPLICANT: Clements, Peter R  
; APPLICANT: Hopwood, John J  
; TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF  
; TITLE OF INVENTION: IDURONATE 2-SULFATASE  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scully, Scott, Murphy & Presser  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,493  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 991,973  
; FILING DATE: 17-DEC-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DiGiglio, Frank S  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 8416Z  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 516-742-4343  
; TELEFAX: 516-742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 510 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-484-493-11

Query Match 17.4%; Score 827.5; DB 1; Length 510;  
Best Local Similarity 27.5%; Pred. No. 6.1e-72;  
Matches 218; Conservative 77; Mismatches 158; Indels 341; Gaps 19;

QY 43 RPNILVLTDDDELGLSMQVMTNTRIMEQGGTHFINAFVTPMCCPSRSSILTQKYVH 102  
DB 3 RPNVLLTDDDELGLSMQVMTNTRIMEQGGTHFINAFVTPMCCPSRSSILTQKYVH 62  
QY 103 NHTYTN--NENCSSPSWQAQHSRTFAVYLS--TGYR--TAFPGKYLYNEYG-----S 151  
DB 63 NHHVNNNTLEGCSKSKWQIQEPNTPPAILRSGMQGYQTFTFFAGKYLYNEYGAPDAGGLE 122  
QY 152 YVPGQKWEVGLLKNSRYNTLCRGVKEKHGSDYKVDLTLDITNDVSP--FRTSKK 209  
DB 123 HVLPGWSYWALEKNSKTYNTLTLSINGKARKHGENYSVDYLTDLVLANVSLDFLDYKSNEE 182  
QY 210 MYPHRPVLMVISHAAPHGPDSPAPQYGRLPNPAQSCHITPSPVYAPNPDKHWIMRYT-GPM 268  
DB 183 -----PFFMMIATPAPHSFWTAAPQYQAFQNVAPRNKFN-IHGTNKHLLIRQAKTGM 236  
QY 269 KPIHNEFTNMLQRKQLTMSVDDSMETIYNNMVLVETGELDNTYIVYVADHGHYIGQGLV 328  
DB 237 TNSIQFLDNAPKRKRWQLLSVDVLVEKLVKRLVFTGELANNYIFYTSNDNGYHTGQFSLP 296

QY 329 KOKSMPIYEDIRVPEYVGPVNVAGCLNPHVILNIDLAPTILDIAGLDI-PADMDGKSIL 387  
DB 297 IDKRQLYEFDIKVPVLLVRGPGIKPNQTSKMLVANIDLGP-ILDIAGYDLANKTQMDGMSLL 355  
QY 388 KLDDTERPVNRFHLKKKVRVVRDLSFLVERGKLLHKRDNDKYDAQEENFLPKYQKRVKDLQC 447  
DB 356 PIL---RGASNL-----TWRSVLYE----- 373  
QY 448 RAEYQTACEQLGQKWQCVEDATGKLKHKCKGPMRLGSSRALSNLVPKYVQGSSEACTCD 507  
DB 374 -----YQEG----- 378  
QY 508 SGDYKLSLAGRRKLLFKKKYKASYVRSRSIRSAIEVDGRVYHVGGLGDAQAQPRNLTKRHW 567  
DB 379 -----RNV----- 382  
QY 568 PGAPEDDDKGGDFSGTGGLPDYSAANPIKVTHRCYILENDTVQCOLDLYKSLQAWKH 627  
DB 383 ----- 382  
QY 628 KLHIDHIEITLQNKILREVVRGHLKKRPEECDCCHKISYHTQHKRGLKHGSSLSHPFRK 687  
DB 383 -----DPTCPSLS----- 390  
QY 688 GLQEKDKVLLREQRKKKLLKRLKRLQNNDTCSMPGLT-CFTHDNQHWQTPPTWTLGPF 746  
DB 391 -----PGVSQCF-----PD 399  
QY 747 CACTSANNTYWCMTINETHNLFCEF--ATGFLVEFDLNTDPVQLMNAVNTLDRDVLN 804  
DB 400 CVCEDAYNTYACVTMTSALNWLQYCEFDQDEVFEVYVNLADPDQITNIAKTIDPELLG 459  
QY 805 QLHVQLMELRSCKG 818  
DB 460 KNNYRLMLQSCSG 473

RESULT 14  
US-08-484-494-11  
; Sequence 11, Application US/08484494  
; Patent No. 5798239  
; GENERAL INFORMATION:  
; APPLICANT: Wilson, Peter J  
; APPLICANT: Morris, Charles P  
; APPLICANT: Anson, Donald S  
; APPLICANT: Occhiodoro, Teresa  
; APPLICANT: Bielsicki, Julie  
; APPLICANT: Clements, Peter R  
; APPLICANT: Hopwood, John J  
; TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF  
; TITLE OF INVENTION: IDURONATE 2-SULFATASE  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scully, Scott, Murphy & Presser  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,494  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 991,973  
; FILING DATE: 17-DEC-1992  
; ATTORNEY/AGENT INFORMATION:

```
; NAME: DiGiglio, Frank S
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8416Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 510 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-484-494-11

Query Match 17.4%; Score 827.5; DB 1; Length 510;
Best Local Similarity 27.5%; Pred. No. 6.1e-72;
Matches 218; Conservative 77; Mismatches 158; Indels 341; Gaps 19;

Qy 43 RPNILVLTDDQDVELGSMQVNMKTRIMEQGGTHFINAFVTPMCCPSRSSILTCKYVH 102
Db 3 RPNVLLTDDQDVELGSMTPKTKKALIGEMGTFSSAYVPSALCCPSRASILTCKYPH 62
Qy 103 NNTYTN--NENCSSPSWQAQHSRTFAVYVLS--TGVR--TAFPGKYLYNEYNG-----S 151
Db 63 NHVNVNTLEGNCSSKSWQKIQEPNTPFALLRSMQGYQTFTFFAGKYLYNEYGAPDAGGLE 122
Qy 152 YVPPGKWEVGLLKNSRFNYTLCRNGVKEKHSYDYLTLTNDVSF--FRTSKK 209
Db 123 HVLPGWSYVALEKSKYNYTLISNGKARKHGENYVDYLTDLVANVSLDFLDYKSNEE 182
Qy 210 MYPHRPVLVISHAAPHGSDSAPQYSRLFPNASQHIPTPSYNAENPKHIMRYT-GPM 268
Db 183 -----PFFMMIATPAPHSWTAPOQKAFQVAFPRNNKFN--IHGTNKGWLIROAKTPM 236
Qy 269 KPIHMEFTNKLQKRLQTLMSVDDSMETIYNNMLVETGELDNTYIYVTADHGYPHIGQGLV 328
Db 237 TNSSIQFLDPAFRKWOTLLSVDDLVEKLVKLEPTGELNNTYIFTSDNGYHTGQSLP 296
Qy 329 KGKSMPEYEDIRVPVVRGPNVEAGCLNPHIVLNTDLAPTLDIAGLDI-PADWOGSKIL 387
Db 297 IDRQYEFEDIKVPVLLVRGPGIKPNQTSKGLVANIDLGP-ILDIAGYDLNKTQMDGMSLL 355
Qy 388 KLLDTRPVRNRFHLKKMKVRDSFLVRGKLLHKRDNDKVDAQENFLPKYQRYKOLCO 447
Db 356 PIL--RGASNL-----TWRSDDLVE----- 373
Qy 448 RASYQTACBQKGQWQCVEDATGKLLHKCKGPMRLGGSRALSNLVPKYQGSGSEACTD 507
Db 374 -----YQEG----- 378
Qy 508 SGDYKLSLAGRRKKLFKKYKASYVRSRSIRSAIEVDGRVYHVLGDAAPRNLTKRHW 567
Db 379 -----RNV----- 382
Qy 568 PGAPEDQDDXGDFSGTGGLPDYSAANPIKUTHRCYILENDTVQCDLDLYKSLQAWKH 627
Db 383 ----- 382
Qy 628 KLHIDHETLQNKINLREVRGLKXKRPEDCHKISYHTQHKRLKRGSSLHPFK 687
Db 383 -----DPTCPSL----- 390
Qy 688 GLQEKVKVLLRSQKRRKKLKLKRLQNNDTCSMPGLT-CPTHQNHQWTAPEFTLGP 746
Db 391 -----PGVSQCF-----PD 399
Qy 747 CACTSANNTYMCWRTINETHNPLFCF--ATGPLEYEDLNTDPYQANVNTLDQVNL 804
Db 400 CVEDATNNTYACVRTWSALWNLQYCEFDQEVFEVYTNLTADPQITNIAKTIDPELLG 459
Qy 805 QLVHQLMELRSCKG 818
```

```
Db 460 KNYRLAMLMQSCSG 473

RESULT 15
US-08-345-212-11
; Sequence 11, Application US/08345212
; Patent No. 5932211
; GENERAL INFORMATION:
; APPLICANT: Mileon, Peter J
; APPLICANT: Morris, Charles P
; APPLICANT: Anson, Donald S
; APPLICANT: Occhiodoro, Teresa
; APPLICANT: Bielicki, Julie
; APPLICANT: Clements, Peter R
; APPLICANT: Hopwood, John J
; TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
; TITLE OF INVENTION: IDURONATE 2-SULFATASE
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/345,212
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 991,973
; FILING DATE: 17-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8416Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 510 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-345-212-11

Query Match 17.4%; Score 827.5; DB 2; Length 510;
Best Local Similarity 27.5%; Pred. No. 6.1e-72;
Matches 218; Conservative 77; Mismatches 158; Indels 341; Gaps 19;

Qy 43 RPNILVLTDDQDVELGSMQVNMKTRIMEQGGTHFINAFVTPMCCPSRSSILTCKYVH 102
Db 3 RPNVLLTDDQDVELGSMTPKTKKALIGEMGTFSSAYVPSALCCPSRASILTCKYPH 62
Qy 103 NNTYTN--NENCSSPSWQAQHSRTFAVYVLS--TGVR--TAFPGKYLYNEYNG-----S 151
Db 63 NHVNVNTLEGNCSSKSWQKIQEPNTPFALLRSMQGYQTFTFFAGKYLYNEYGAPDAGGLE 122
Qy 152 YVPPGKWEVGLLKNSRFNYTLCRNGVKEKHSYDYLTLTNDVSF--FRTSKK 209
Db 123 HVLPGWSYVALEKSKYNYTLISNGKARKHGENYVDYLTDLVANVSLDFLDYKSNEE 182
Qy 210 MYPHRPVLVISHAAPHGSDSAPQYSRLFPNASQHIPTPSYNAENPKHIMRYT-GPM 268
```

Db 183 -----PFFMMIATPAPHSEPWTAAPQYQKAFQNVFAPRNKNFN-IHGTNKHMLIRQAKTPM 236  
QY 269 KP1HMEFTNMLQKRLOTLSVDDSDNETIYNMLVTGELDNITYIUYTADHGHIHQGGLV 328  
Db 237 TNSSIQFLDNAFKRWQTLSSVDDLVEKLVKRLBFTGELNNTYIIFYTSDNGYHTGQFSLP 296  
QY 329 KGKSMPEYFDIRVFPVVRGPNVEAGCLNPHIVLNIDLAPTILDIAGLDI-PADMDGKSIL 387  
Db 297 IDKRLYEFDIKVPLLVVRGPIKPNQTSKMLVANIDLGP-ILDIAGYDLNKTQMDGWSLL 355  
QY 388 KLLDTERPVRNFKKKMRVWRDSPLVERKLLHKRDNKDVAQBEENFLPKYQVRKDLQ 447  
Db 356 PIL---RGASNL-----TWRSDDLVE----- 373  
QY 448 RAEYQTAQELQKQWQVEDATGKXLHKCKGPMRLGGSRLSNLVPKYQGSEACTCD 507  
Db 374 -----YQSEG----- 378  
QY 508 SGDYKLSLAGRRKKLPKKYKASYRSTRSIRVAIEVDGRVYHVGIDAAQPRNLTKRHW 567  
Db 379 -----RNV----- 382  
QY 568 PGAPEDQDDKDGDFSGTGLPDYSAANPIKVTHRCYILENDTVOCDDLKLYSLQAWKDH 627  
Db 383 ----- 382  
QY 628 KLIHDHIEITLQNKIKNLREVRGHLKKGPECCDCHKISYHTOHKGRKLRGSSLHPFRK 687  
Db 383 -----DPTCPSLS----- 390  
QY 688 GLQEKDKVMLLREQKRRKKLRLKRLQNNDTCSMPLGT-CFTHDNQHWQTAPFWTLGPF 746  
Db 391 -----PGVSQCF-----PD 399  
QY 747 CACTSANNTYMCWRTINETHNFCEP--ATGFLEYFDLNTDPYQLANAVNTLDRDVLN 804  
Db 400 CVCEDAYNTYACVRTMSALWNLQYCEFDQEVFEVYNLTADPDQITNIAKTIDPELLG 459  
QY 805 QLVQVLMELRSCKG 818  
Db 460 KMYRLMLQSCSG 473

Search completed: October 15, 2003, 13:10:16  
Job time : 24 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2003, 13:09:25 ; Search time 77 Seconds

(without alignments)  
1820.542 Million cell updates/sec

Title: US-09-970-287-2

Perfect score: 4751

Sequence: 1 MGPPSLVLCGLSATVFLSLG.....PEMKRPSKSLGQLWEGWEG 870

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 600653 seqs, 161128416 residues

Total number of hits satisfying chosen parameters: 600653

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4751	100.0	870	12	US-09-970-287-2
2	4751	100.0	885	12	US-09-970-287-12
3	4746	99.9	870	12	US-10-182-951-5
4	4746	99.9	870	12	US-10-025-966A-6
5	4746	99.9	870	12	US-10-025-966A-15
6	4746	99.9	870	12	US-10-265-071-6
7	4746	99.9	870	12	US-10-265-071-15
8	4719.5	99.3	867	9	US-09-989-722-84
9	4719.5	99.3	867	9	US-09-989-723-84
10	4719.5	99.3	867	9	US-09-989-279-84
11	4719.5	99.3	867	9	US-09-989-727-84
12	4719.5	99.3	867	10	US-09-989-731-84
13	4719.5	99.3	867	10	US-09-989-732-84
14	4719.5	99.3	867	10	US-09-991-073-84
15	4719.5	99.3	867	10	US-09-990-442-84

16	4719.5	99.3	867	10	US-09-991-163-84	Sequence 84, Appl
17	4719.5	99.3	867	10	US-09-993-604-84	Sequence 84, Appl
18	4719.5	99.3	867	10	US-09-990-456-84	Sequence 84, Appl
19	4719.5	99.3	867	10	US-09-989-721-84	Sequence 84, Appl
20	4719.5	99.3	867	10	US-09-992-598-84	Sequence 84, Appl
21	4719.5	99.3	867	10	US-09-989-293A-84	Sequence 84, Appl
22	4719.5	99.3	867	10	US-09-989-735-84	Sequence 84, Appl
23	4719.5	99.3	867	10	US-09-990-444-84	Sequence 84, Appl
24	4719.5	99.3	867	10	US-09-991-181-84	Sequence 84, Appl
25	4719.5	99.3	867	10	US-09-989-730-84	Sequence 84, Appl
26	4719.5	99.3	867	10	US-09-990-436-84	Sequence 84, Appl
27	4719.5	99.3	867	10	US-09-993-687-84	Sequence 84, Appl
28	4719.5	99.3	867	11	US-09-989-734-84	Sequence 84, Appl
29	4719.5	99.3	867	11	US-09-997-653-84	Sequence 84, Appl
30	4719.5	99.3	867	11	US-09-993-667-84	Sequence 84, Appl
31	4719.5	99.3	867	11	US-09-997-428-84	Sequence 84, Appl
32	4719.5	99.3	867	11	US-09-997-666-84	Sequence 84, Appl
33	4719.5	99.3	867	11	US-09-990-438-84	Sequence 84, Appl
34	4719.5	99.3	867	11	US-09-990-562-84	Sequence 84, Appl
35	4719.5	99.3	867	11	US-09-990-711-84	Sequence 84, Appl
36	4719.5	99.3	867	11	US-09-989-726-84	Sequence 84, Appl
37	4719.5	99.3	867	11	US-09-998-156-84	Sequence 84, Appl
38	4719.5	99.3	867	11	US-09-990-437-84	Sequence 84, Appl
39	4719.5	99.3	867	11	US-09-991-157-84	Sequence 84, Appl
40	4719.5	99.3	867	11	US-09-997-514-84	Sequence 84, Appl
41	4719.5	99.3	867	11	US-09-997-573-84	Sequence 84, Appl
42	4719.5	99.3	867	11	US-09-991-172-84	Sequence 84, Appl
43	4719.5	99.3	867	11	US-09-990-726-84	Sequence 84, Appl
44	4719.5	99.3	867	11	US-09-997-559-84	Sequence 84, Appl
45	4719.5	99.3	867	11	US-09-997-601-84	Sequence 84, Appl

## ALIGNMENTS

RESULT 1  
US-09-970-287-2  
; Sequence 2, Application US/09970287  
; Publication No. US20030166283A1  
; GENERAL INFORMATION:  
; APPLICANT: GLUCKSMANN, Maria A.  
; APPLICANT: MEYERS, Rachel  
; APPLICANT: KAPILLER-LIBERMANN, Rosana  
; APPLICANT: SILOS-SANTIAGO, Inmaculada  
; TITLE OF INVENTION: 22437, A NOVEL HUMAN SULFATASE AND USES THEREFOR  
; FILE REFERENCE: 10147-61U1  
; CURRENT APPLICATION NUMBER: US/09/970,287  
; PRIOR FILING DATE: 2001-10-03  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 870  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-970-287-2

Query Match	100.0%	Score 4751;	DB 12;	Length 870;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 870;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MGPPSLVLCGLSATVFLSLGSSAFLSHRLKGFQDRNRIRNIIILVLTDDQVELGS	60	
Db	1	MGPPSLVLCGLSATVFLSLGSSAFLSHRLKGFQDRNRIRNIIILVLTDDQVELGS	60	
Qy	61	MQVMKTRIMEQGTTHFINAFVTTMCCPSRSSILTKGYVHNHTYTNNECSSPSWQA	120	
Db	61	MQVMKTRIMEQGTTHFINAFVTTMCCPSRSSILTKGYVHNHTYTNNECSSPSWQA	120	
Qy	121	QHESTRFAVILNSTGYRTAFPGKYLNEYNGSVYPPGKWWGLLNKSRFYNYTLCRNGVK	180	
Db	121	QHESTRFAVILNSTGYRTAFPGKYLNEYNGSVYPPGKWWGLLNKSRFYNYTLCRNGVK	180	

QY 181 EKHSYKDYKYLTLITNDVSFPRTSKQYPRPVLVMIASHAHPGSDSAPQYSRLFP 240  
 DB 181 EKHSYKDYKYLTLITNDVSFPRTSKQYPRPVLVMIASHAHPGSDSAPQYSRLFP 240  
 QY 241 NASQHIITPSYNAFNPDKHIMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300  
 DB 241 NASQHIITPSYNAFNPDKHIMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300  
 QY 301 LVETGELDNITYVTADHGYHIGQFGLVKGKSMPEYEDIRVPPYVRGPNVEAGCLNPHIV 360  
 DB 301 LVETGELDNITYVTADHGYHIGQFGLVKGKSMPEYEDIRVPPYVRGPNVEAGCLNPHIV 360  
 QY 361 LNIDLAPTILDIAGLDIPADMDGKSIILKLDTERPVRNPHLKKMRVWRDSFLVERGKLL 420  
 DB 361 LNIDLAPTILDIAGLDIPADMDGKSIILKLDTERPVRNPHLKKMRVWRDSFLVERGKLL 420  
 QY 421 HKRDNKVDQAEENFLPKYQRYKDLQORAEYQACBOLGQKMQCVEDATGKLLKHKCKGP 480  
 DB 421 HKRDNKVDQAEENFLPKYQRYKDLQORAEYQACBOLGQKMQCVEDATGKLLKHKCKGP 480  
 QY 481 MRLGGSRAALSNLVPKYGGSEACTCDSGDYKLSLAGRRKGLFKKYYKASYRSTRSIRSV 540  
 DB 481 MRLGGSRAALSNLVPKYGGSEACTCDSGDYKLSLAGRRKGLFKKYYKASYRSTRSIRSV 540  
 QY 541 AIEVDGRVYHVLGDAAPRNLTKRHWPGAPEDODDKOGDFSGTGGLPDYSAANPIKVT 600  
 DB 541 AIEVDGRVYHVLGDAAPRNLTKRHWPGAPEDODDKOGDFSGTGGLPDYSAANPIKVT 600  
 QY 601 HRCYILENDTVQCDLDLYKSLQAWKHKLHIDHEIETLQNKIKNLRVGRHLKKRPEEC 660  
 DB 601 HRCYILENDTVQCDLDLYKSLQAWKHKLHIDHEIETLQNKIKNLRVGRHLKKRPEEC 660  
 QY 661 DCHKISYHTQHKHGRGSSLHPFRKGLQEKDKVWLLREQRKKKLLKRLQNNDT 720  
 DB 661 DCHKISYHTQHKHGRGSSLHPFRKGLQEKDKVWLLREQRKKKLLKRLQNNDT 720  
 QY 721 SMPGLTCFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNFCEPATGFLE 780  
 DB 721 SMPGLTCFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNFCEPATGFLE 780  
 QY 781 YFDLNTDYPQLMNAVNTLDRDLVNLQHLVQLMELRSCGYKQCNPRTRNMDLGLKDGGSYE 840  
 DB 781 YFDLNTDYPQLMNAVNTLDRDLVNLQHLVQLMELRSCGYKQCNPRTRNMDLGLKDGGSYE 840  
 QY 841 QYRQFQRKWPCKRPPSSKSLGQWEGWEG 870  
 DB 841 QYRQFQRKWPCKRPPSSKSLGQWEGWEG 870

## RESULT 2

US-09-970-287-12  
 ; Sequence 12, Application US/09970287  
 ; Publication No. US20030166283A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MEYERS, Rachel  
 ; APPLICANT: GLUCKSMANN, Maria A.  
 ; APPLICANT: MEYERS, Rachel  
 ; APPLICANT: KAPELLER-LIBERMANN, Rosana  
 ; APPLICANT: SILOS-SANTIAGO, Inmaculada  
 ; TITLE OF INVENTION: 22437, A NOVEL HUMAN SULFATASE AND USES THEREFOR  
 ; FILE REFERENCE: 10147-6101  
 ; CURRENT APPLICATION NUMBER: US/09/970,287  
 ; CURRENT FILING DATE: 2001-10-03  
 ; PRIOR APPLICATION NUMBER: US 60/257,082  
 ; PRIOR FILING DATE: 2000-12-21  
 ; NUMBER OF SEQ ID NOS: 12  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 12  
 ; LENGTH: 885  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-970-287-12

Query Match 100.0%; Score 4751; DB 12; Length 885;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MGPPSLVLCCLLSATVPSLLGGSSAFSLSHRLKGRFQDRRNIRPNIILVLTDDQDVELGS 60  
 DB 16 MGPPSLVLCCLLSATVPSLLGGSSAFSLSHRLKGRFQDRRNIRPNIILVLTDDQDVELGS 75  
 QY 61 MOVAKTRIMEQGTTHPINAFTVTPMCCPSRSISLTGKYVHNHTYNNENCSSPSWQA 120  
 DB 76 MOVAKTRIMEQGTTHPINAFTVTPMCCPSRSISLTGKYVHNHTYNNENCSSPSWQA 135  
 QY 121 QHESRTFAVILNSTGYRTAFPGKYLNEVNGSYVPPGKWEVGLLKNRFFNYTLCRNGVK 180  
 DB 136 QHESRTFAVILNSTGYRTAFPGKYLNEVNGSYVPPGKWEVGLLKNRFFNYTLCRNGVK 195  
 QY 181 EKHSYKDYKYLTLITNDVSFPRTSKQYPRPVLVMIASHAHPGSDSAPQYSRLFP 240  
 DB 196 EKHSYKDYKYLTLITNDVSFPRTSKQYPRPVLVMIASHAHPGSDSAPQYSRLFP 255  
 QY 241 NASQHIITPSYNAFNPDKHIMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300  
 DB 256 NASQHIITPSYNAFNPDKHIMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 315  
 QY 301 LVETGELDNITYVTADHGYHIGQFGLVKGKSMPEYEDIRVPPYVRGPNVEAGCLNPHIV 360  
 DB 316 LVETGELDNITYVTADHGYHIGQFGLVKGKSMPEYEDIRVPPYVRGPNVEAGCLNPHIV 375  
 QY 361 LNIDLAPTILDIAGLDIPADMDGKSIILKLDTERPVRNPHLKKMRVWRDSFLVERGKLL 420  
 DB 376 LNIDLAPTILDIAGLDIPADMDGKSIILKLDTERPVRNPHLKKMRVWRDSFLVERGKLL 435  
 QY 421 HKRDNKVDQAEENFLPKYQRYKDLQORAEYQACBOLGQKMQCVEDATGKLLKHKCKGP 480  
 DB 436 HKRDNKVDQAEENFLPKYQRYKDLQORAEYQACBOLGQKMQCVEDATGKLLKHKCKGP 495  
 QY 481 MRLGGSRAALSNLVPKYGGSEACTCDSGDYKLSLAGRRKGLFKKYYKASYRSTRSIRSV 540  
 DB 496 MRLGGSRAALSNLVPKYGGSEACTCDSGDYKLSLAGRRKGLFKKYYKASYRSTRSIRSV 555  
 QY 541 AIEVDGRVYHVLGDAAPRNLTKRHWPGAPEDODDKOGDFSGTGGLPDYSAANPIKVT 600  
 DB 556 AIEVDGRVYHVLGDAAPRNLTKRHWPGAPEDODDKOGDFSGTGGLPDYSAANPIKVT 615  
 QY 601 HRCYILENDTVQCDLDLYKSLQAWKHKLHIDHEIETLQNKIKNLRVGRHLKKRPEEC 660  
 DB 616 HRCYILENDTVQCDLDLYKSLQAWKHKLHIDHEIETLQNKIKNLRVGRHLKKRPEEC 675  
 QY 661 DCHKISYHTQHKHGRGSSLHPFRKGLQEKDKVWLLREQRKKKLLKRLQNNDT 720  
 DB 676 DCHKISYHTQHKHGRGSSLHPFRKGLQEKDKVWLLREQRKKKLLKRLQNNDT 735  
 QY 721 SMPGLTCFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNFCEPATGFLE 780  
 DB 736 SMPGLTCFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNFCEPATGFLE 795  
 QY 781 YFDLNTDYPQLMNAVNTLDRDLVNLQHLVQLMELRSCGYKQCNPRTRNMDLGLKDGGSYE 840  
 DB 796 YFDLNTDYPQLMNAVNTLDRDLVNLQHLVQLMELRSCGYKQCNPRTRNMDLGLKDGGSYE 855  
 QY 841 QYRQFQRKWPCKRPPSSKSLGQWEGWEG 870  
 DB 856 QYRQFQRKWPCKRPPSSKSLGQWEGWEG 885

## RESULT 3

US-10-182-951-5  
 ; Sequence 5, Application US/10182951  
 ; Publication No. US20030138895A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: INCYTE GENOMICS, INC.  
 ; APPLICANT: TANG, Y. Tom  
 ; APPLICANT: BAUGHN, Mariah R.

```
; APPLICANT: YAO, Monique G.
; APPLICANT: BANDMAN, Olga
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: LAL, Preeti
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: RING, Huijun Z.
; APPLICANT: SHIH, Leo L.
; APPLICANT: YANG, Junming
; APPLICANT: POLICKY, Jennifer L.
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
; CURRENT APPLICATION NUMBER: US/10/182,951
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/181,856; 10/183,684; 60/185,141; 60/186,818; 60/188,345;
; PRIOR FILING DATE: 2000-02-11; 2000-02-17; 2000-02-25; 2000-03-03; 2000-03-09; 2000-0
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 870
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030138895A1 1558210CD1
US-10-182-951-5

Query Match          99.9%; Score 4746; DB 12; Length 870;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGPPSLVCLLSATVFSLLGGSSAFSLSHRLKGRFDRNRNIRPNILVLTDQDVELGS 60
Db 1 MGPPSLVCLLSATVFSLLGGSSAFSLSHRLKGRFDRNRNIRPNILVLTDQDVELGS 60
Qy 61 MQVMKTRRIMEQGGTHFINAVTTTMCPCSSRSILTKGYVHHNTYTNNECSSPSWQA 120
Db 61 MQVMKTRRIMEQGGTHFINAVTTTMCPCSSRSILTKGYVHHNTYTNNECSSPSWQA 120
Qy 121 QHESRTFAVLNSTGYRTAFPGKYLYNEYSYVPPGKWEVGLLKNRSFYNYTLCRNGVK 180
Db 121 QHESRTFAVLNSTGYRTAFPGKYLYNEYSYVPPGKWEVGLLKNRSFYNYTLCRNGVK 180
Qy 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKQMPYHPRLVLAISHAAPHGSDSAPQYSRLFP 240
Db 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKQMPYHPRLVLAISHAAPHGSDSAPQYSRLFP 240
Qy 241 NASOHITPSYNYAPNDKHWIMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
Db 241 NASOHITPSYNYAPNDKHWIMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
Qy 301 LVETGELONTYIVYTADHGHIQFGLVKGSMPYEFDIRVPPYVRGPNVEAGCLNPHIV 360
Db 301 LVETGELONTYIVYTADHGHIQFGLVKGSMPYEFDIRVPPYVRGPNVEAGCLNPHIV 360
Qy 361 LNIIDLAPTILDIAGLDIPADMDGKSLKLDTERPVNRFLKXKORVWRDSEFLVERGKLL 420
Db 361 LNIIDLAPTILDIAGLDIPADMDGKSLKLDTERPVNRFLKXKORVWRDSEFLVERGKLL 420
Qy 421 HCRDNKVDQAEENFLPKYQVVKDLQARAEYQACBQLGQKWCQVEDATGKLKHKCKGP 480
Db 421 HCRDNKVDQAEENFLPKYQVVKDLQARAEYQACBQLGQKWCQVEDATGKLKHKCKGP 480
Qy 481 MRLGSSRALSNLVPKYQGSEACTCDSDGYKLSLAGRKKLPKXYKASVYRSRSIRSV 540
Db 481 MRLGSSRALSNLVPKYQGSEACTCDSDGYKLSLAGRKKLPKXYKASVYRSRSIRSV 540
Qy 541 AIEVDGRVYHVLGDAQAQRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
Db 541 AIEVDGRVYHVLGDAQAQRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
Qy 601 HRCVILENDTVQCDLDLYKSLQAWKHKHIHDIETLQNKIKNREVRCHLKKRPREEC 660
Db 601 HRCVILENDTVQCDLDLYKSLQAWKHKHIHDIETLQNKIKNREVRCHLKKRPREEC 660

RESULT 4
US-10-025-966A-6
; Sequence 6, Application US/10025966A
; Publication No. US20030148920A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Steven D.
; APPLICANT: Palmeri, Diana
; APPLICANT: Stefan, Hennerich
; TITLE OF INVENTION: Sulfatases and methods of use
; FILE REFERENCE: UCAL230
; CURRENT APPLICATION NUMBER: US/10/025,966A
; PRIOR APPLICATION NUMBER: 2002-12-21
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 60/267,831
; PRIOR FILING DATE: 2001-09-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 870
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-025-966A-6

Query Match          99.9%; Score 4746; DB 12; Length 870;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGPPSLVCLLSATVFSLLGGSSAFSLSHRLKGRFDRNRNIRPNILVLTDQDVELGS 60
Db 1 MGPPSLVCLLSATVFSLLGGSSAFSLSHRLKGRFDRNRNIRPNILVLTDQDVELGS 60
Qy 61 MQVMKTRRIMEQGGTHFINAVTTTMCPCSSRSILTKGYVHHNTYTNNECSSPSWQA 120
Db 61 MQVMKTRRIMEQGGTHFINAVTTTMCPCSSRSILTKGYVHHNTYTNNECSSPSWQA 120
Qy 121 QHESRTFAVLNSTGYRTAFPGKYLYNEYSYVPPGKWEVGLLKNRSFYNYTLCRNGVK 180
Db 121 QHESRTFAVLNSTGYRTAFPGKYLYNEYSYVPPGKWEVGLLKNRSFYNYTLCRNGVK 180
Qy 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKQMPYHPRLVLAISHAAPHGSDSAPQYSRLFP 240
Db 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKQMPYHPRLVLAISHAAPHGSDSAPQYSRLFP 240
Qy 241 NASOHITPSYNYAPNDKHWIMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
Db 241 NASOHITPSYNYAPNDKHWIMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
Qy 301 LVETGELONTYIVYTADHGHIQFGLVKGSMPYEFDIRVPPYVRGPNVEAGCLNPHIV 360
Db 301 LVETGELONTYIVYTADHGHIQFGLVKGSMPYEFDIRVPPYVRGPNVEAGCLNPHIV 360
Qy 361 LNIIDLAPTILDIAGLDIPADMDGKSLKLDTERPVNRFLKXKORVWRDSEFLVERGKLL 420
Db 361 LNIIDLAPTILDIAGLDIPADMDGKSLKLDTERPVNRFLKXKORVWRDSEFLVERGKLL 420
```

```

Db 361 LNTDLAPTLDIAGLDIPADMDGKSLKLLDTERPVRNPHLKKORVWRDSFLVERGKLL 420
Qy 421 HKRDNKVDQAQENFLPKYQRYKDLQRAEYQTAQCEQLGQKQWQCVEDATGKLLHKCKGP 480
Db 421 HKRDNKVDQAQENFLPKYQRYKDLQRAEYQTAQCEQLGQKQWQCVEDATGKLLHKCKGP 480
Qy 481 MRGGSRALSNLVPKYGGSEACTCDSGDYKLSLAGRRKLLPKKYKASYVRSRSIRSV 540
Db 481 MRGGSRALSNLVPKYGGSEACTCDSGDYKLSLAGRRKLLPKKYKASYVRSRSIRSV 540
Qy 541 AIEVDGRVHVGLGDAAPRNLTKEHWPFGAPEDQDDKQGGDFSGTGGLPDYSAANPIKVT 600
Db 541 AIEVDGRVHVGLGDAAPRNLTKEHWPFGAPEDQDDKQGGDFSGTGGLPDYSAANPIKVT 600
Qy 601 HRCYILENDTVQCDLQYKSLQAKHKLHIDHEITLQNKIKMLREVRGHLKKGP 660
Db 601 HRCYILENDTVQCDLQYKSLQAKHKLHIDHEITLQNKIKMLREVRGHLKKGP 660
Qy 661 DCHKISYHTQHKGRLLKRGSSLHPFRKGLQKQKVMLLREQRKKLLKLLKELQNNDT 720
Db 661 DCHKISYHTQHKGRLLKRGSSLHPFRKGLQKQKVMLLREQRKKLLKLLKELQNNDT 720
Qy 721 SMPGLTCFTHDNQHQWQTAFFTLGPFCACTSANNTYWCMTINETHNFCEPATGFL 780
Db 721 SMPGLTCFTHDNQHQWQTAFFTLGPFCACTSANNTYWCMTINETHNFCEPATGFL 780
Qy 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCGKYQKQCNPRTRNMDLGLKDGGSYE 840
Db 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCGKYQKQCNPRTRNMDLGLKDGGSYE 840
Qy 841 QYRQFORRWKPEMKRPFSSKSLGQLWEGWEG 870
Db 841 QYRQFORRWKPEMKRPFSSKSLGQLWEGWEG 870

RESULT 5
US-10-025-966A-15
; Sequence 15, Application US/10025966A
; Publication No. US20030148920A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Steven D.
; APPLICANT: Palmeri, Diana
; APPLICANT: Stefan, Hemmerich
; TITLE OF INVENTION: Sulfatases and methods of use
; TITLE OF INVENTION: thereof
; FILE REFERENCE: UCAL230
; CURRENT APPLICATION NUMBER: US/10/025,966A
; CURRENT FILING DATE: 2002-12-21
; PRIOR APPLICATION NUMBER: 60/258,577
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 60/267,831
; PRIOR FILING DATE: 2001-09-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 870
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-025-966A-15

Query Match 99.9%; Score 4746; DB 12; Length 870;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGPPSLVLCLLSATVPSLGGSSAFLSHRLKGRFQDRDRIRNIIVLVTDQDVELGS 60
Db 1 MGPPSLVLCLLSATVPSLGGSSAFLSHRLKGRFQDRDRIRNIIVLVTDQDVELGS 60
Qy 61 MQMNTKTRIMEGGGTHFINAVPTTMCPCPSRSSLTGKYVHNNTYTNKNCSSPSWQA 120
Db 61 MQMNTKTRIMEGGGTHFINAVPTTMCPCPSRSSLTGKYVHNNTYTNKNCSSPSWQA 120
Qy 121 QHSRTFAVYLNSTGYRTAFPGKYLYNEYSYVPPGKEMVGLLKNRFFNYTLCRNGVK 180

```

```

Db 121 QHSRTFAVYLNSTGYRTAFPGKYLYNEYSYVPPGKEMVGLLKNRFFNYTLCRNGVK 180
Qy 181 EXGSDYSKDYLTDLITDNDVSVPFRTSKQMPHPRVPLMVI SHAAAPHGPEDSAPOYSRLFP 240
Db 181 EXGSDYSKDYLTDLITDNDVSVPFRTSKQMPHPRVPLMVI SHAAAPHGPEDSAPOYSRLFP 240
Qy 241 NASQHTTPSYNTAPNDKINIMRYTGPMPKPIMHEFTNMLQRRLQTLMSVDDSHETTINM 300
Db 241 NASQHTTPSYNTAPNDKINIMRYTGPMPKPIMHEFTNMLQRRLQTLMSVDDSHETTINM 300
Qy 301 LVETGELDNTYIVYADHGTHIGQFGLVKGKSMPEFDIRVPPFYVRGPNVEAGCLNPHIV 360
Db 301 LVETGELDNTYIVYADHGTHIGQFGLVKGKSMPEFDIRVPPFYVRGPNVEAGCLNPHIV 360
Qy 361 LNTDLAPTLDIAGLDIPADMDGKSLKLLDTERPVRNPHLKKORVWRDSFLVERGKLL 420
Db 361 LNTDLAPTLDIAGLDIPADMDGKSLKLLDTERPVRNPHLKKORVWRDSFLVERGKLL 420
Qy 421 HKRDNKVDQAQENFLPKYQRYKDLQRAEYQTAQCEQLGQKQWQCVEDATGKLLHKCKGP 480
Db 421 HKRDNKVDQAQENFLPKYQRYKDLQRAEYQTAQCEQLGQKQWQCVEDATGKLLHKCKGP 480
Qy 481 MRGGSRALSNLVPKYGGSEACTCDSGDYKLSLAGRRKLLPKKYKASYVRSRSIRSV 540
Db 481 MRGGSRALSNLVPKYGGSEACTCDSGDYKLSLAGRRKLLPKKYKASYVRSRSIRSV 540
Qy 541 AIEVDGRVHVGLGDAAPRNLTKEHWPFGAPEDQDDKQGGDFSGTGGLPDYSAANPIKVT 600
Db 541 AIEVDGRVHVGLGDAAPRNLTKEHWPFGAPEDQDDKQGGDFSGTGGLPDYSAANPIKVT 600
Qy 601 HRCYILENDTVQCDLQYKSLQAKHKLHIDHEITLQNKIKMLREVRGHLKKGP 660
Db 601 HRCYILENDTVQCDLQYKSLQAKHKLHIDHEITLQNKIKMLREVRGHLKKGP 660
Qy 661 DCHKISYHTQHKGRLLKRGSSLHPFRKGLQKQKVMLLREQRKKLLKLLKELQNNDT 720
Db 661 DCHKISYHTQHKGRLLKRGSSLHPFRKGLQKQKVMLLREQRKKLLKLLKELQNNDT 720
Qy 721 SMPGLTCFTHDNQHQWQTAFFTLGPFCACTSANNTYWCMTINETHNFCEPATGFL 780
Db 721 SMPGLTCFTHDNQHQWQTAFFTLGPFCACTSANNTYWCMTINETHNFCEPATGFL 780
Qy 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCGKYQKQCNPRTRNMDLGLKDGGSYE 840
Db 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCGKYQKQCNPRTRNMDLGLKDGGSYE 840
Qy 841 QYRQFORRWKPEMKRPFSSKSLGQLWEGWEG 870
Db 841 QYRQFORRWKPEMKRPFSSKSLGQLWEGWEG 870

RESULT 6
US-10-265-071-6
; Sequence 6, Application US/10265071
; Publication No. US20030147875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Steven D.
; APPLICANT: Hemmerich, Stefan
; APPLICANT: Tomita, Megumi
; TITLE OF INVENTION: Sulfatases and methods of use thereof
; FILE REFERENCE: UCAL-230CIP
; CURRENT APPLICATION NUMBER: US/10/265,071
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,577
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 60/267,831
; PRIOR FILING DATE: 2001-09-02
; PRIOR APPLICATION NUMBER: 10/025,966
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6

```



```

; LENGTH: 870
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-265-071-6

Query Match      99.9%; Score 4746; DB 12; Length 870;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGPPSLVCLLSATVFSLLGSSAFSLSHRLKGRFQDRNRIRPNILVLTDDQDELGS 60
Db 1 MGPPSLVCLLSATVFSLLGSSAFSLSHRLKGRFQDRNRIRPNILVLTDDQDELGS 60
Qy 61 MQVMNTRIMEGGTHFNAFVTPMCCPSRSSILTKGYVHNHTYNNENCSSPSQA 120
Db 61 MQVMNTRIMEGGTHFNAFVTPMCCPSRSSILTKGYVHNHTYNNENCSSPSQA 120
Qy 121 QHESRTFAVTLNSTGRTAFPGKYLNEYNGSVVPPGKEMWGLKNSRFNYTLCRNGVK 180
Db 121 QHESRTFAVTLNSTGRTAFPGKYLNEYNGSVVPPGKEMWGLKNSRFNYTLCRNGVK 180
Qy 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKOMYPHRLVWVISHAAPHGDESAPOYSRLFP 240
Db 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKOMYPHRLVWVISHAAPHGDESAPOYSRLFP 240
Qy 241 NASOHITPSYNYAPNPKHWMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
Db 241 NASOHITPSYNYAPNPKHWMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
Qy 301 LVETGELDNITYIVYTADHGYPHIGQFGLVKGSMPEYFDIRVFPYVGRGNVEAGCLNPHIV 360
Db 301 LVETGELDNITYIVYTADHGYPHIGQFGLVKGSMPEYFDIRVFPYVGRGNVEAGCLNPHIV 360
Qy 361 LNIDLAPTILDIAGLDIPADMKGSIKLKLDTERPVRNPHLKKKRVWRDSFLVERGKLL 420
Db 361 LNIDLAPTILDIAGLDIPADMKGSIKLKLDTERPVRNPHLKKKRVWRDSFLVERGKLL 420
Qy 421 HKRDNDKVDAAQENFLPKYQVNDLCORAEYQACBQLGQKWCVEDATGKLKHKCKGP 480
Db 421 HKRDNDKVDAAQENFLPKYQVNDLCORAEYQACBQLGQKWCVEDATGKLKHKCKGP 480
Qy 481 MRGGSRALSNLVPKYGGSEACTCDSDYKLSLAGRRKKLFKKYKASVRSRSIRSV 540
Db 481 MRGGSRALSNLVPKYGGSEACTCDSDYKLSLAGRRKKLFKKYKASVRSRSIRSV 540
Qy 541 AIEVDGRVYHVLGDAQAQPRNLTKRHWPCAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
Db 541 AIEVDGRVYHVLGDAQAQPRNLTKRHWPCAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
Qy 601 HRCYIILENDTVQCDDLDLYKSLQAWKHKLHIHDEIETLQNKIKNLRVGRGHLKKRPEEC 660
Db 601 HRCYIILENDTVQCDDLDLYKSLQAWKHKLHIHDEIETLQNKIKNLRVGRGHLKKRPEEC 660
Qy 721 SMPGLTCFTHDNHWTAPFWTLGPPCACTSANNTYWCMTINETHNPLFCBATGPLE 780
Db 721 SMPGLTCFTHDNHWTAPFWTLGPPCACTSANNTYWCMTINETHNPLFCBATGPLE 780
Qy 781 YPDLNTDPYQLMNAVTLDRDVLNQLHVLQVLMELRSCKYKOCNPRTRNMDIQLKDGGSYE 840
Db 781 YPDLNTDPYQLMNAVTLDRDVLNQLHVLQVLMELRSCKYKOCNPRTRNMDIQLKDGGSYE 840
Qy 841 QYRQFORRWKPEMKRPPSSSLGQLMWEGWEG 870
Db 841 QYRQFORRWKPEMKRPPSSSLGQLMWEGWEG 870

```

RESULT 7  
JS-10-265-071-15  
, Sequence 15, Application US/10265071

```

; Publication No. US20030147875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Steven D.
; APPLICANT: Hemmerich, Stefan
; APPLICANT: Tomita, Megumi
; TITLE OF INVENTION: Sulfatases and methods of use thereof
; FILE REFERENCE: UCAL-230CIP
; CURRENT APPLICATION NUMBER: US/10/265,071
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,577
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 60/267,831
; PRIOR FILING DATE: 2001-09-02
; PRIOR APPLICATION NUMBER: 10/025,966
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 870
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-265-071-15

Query Match      99.9%; Score 4746; DB 12; Length 870;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGPPSLVCLLSATVFSLLGSSAFSLSHRLKGRFQDRNRIRPNILVLTDDQDELGS 60
Db 1 MGPPSLVCLLSATVFSLLGSSAFSLSHRLKGRFQDRNRIRPNILVLTDDQDELGS 60
Qy 61 MQVMNTRIMEGGTHFNAFVTPMCCPSRSSILTKGYVHNHTYNNENCSSPSQA 120
Db 61 MQVMNTRIMEGGTHFNAFVTPMCCPSRSSILTKGYVHNHTYNNENCSSPSQA 120
Qy 121 QHESRTFAVTLNSTGRTAFPGKYLNEYNGSVVPPGKEMWGLKNSRFNYTLCRNGVK 180
Db 121 QHESRTFAVTLNSTGRTAFPGKYLNEYNGSVVPPGKEMWGLKNSRFNYTLCRNGVK 180
Qy 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKOMYPHRLVWVISHAAPHGDESAPOYSRLFP 240
Db 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKOMYPHRLVWVISHAAPHGDESAPOYSRLFP 240
Qy 241 NASOHITPSYNYAPNPKHWMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
Db 241 NASOHITPSYNYAPNPKHWMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
Qy 301 LVETGELDNITYIVYTADHGYPHIGQFGLVKGSMPEYFDIRVFPYVGRGNVEAGCLNPHIV 360
Db 301 LVETGELDNITYIVYTADHGYPHIGQFGLVKGSMPEYFDIRVFPYVGRGNVEAGCLNPHIV 360
Qy 361 LNIDLAPTILDIAGLDIPADMKGSIKLKLDTERPVRNPHLKKKRVWRDSFLVERGKLL 420
Db 361 LNIDLAPTILDIAGLDIPADMKGSIKLKLDTERPVRNPHLKKKRVWRDSFLVERGKLL 420
Qy 421 HKRDNDKVDAAQENFLPKYQVNDLCORAEYQACBQLGQKWCVEDATGKLKHKCKGP 480
Db 421 HKRDNDKVDAAQENFLPKYQVNDLCORAEYQACBQLGQKWCVEDATGKLKHKCKGP 480
Qy 481 MRGGSRALSNLVPKYGGSEACTCDSDYKLSLAGRRKKLFKKYKASVRSRSIRSV 540
Db 481 MRGGSRALSNLVPKYGGSEACTCDSDYKLSLAGRRKKLFKKYKASVRSRSIRSV 540
Qy 541 AIEVDGRVYHVLGDAQAQPRNLTKRHWPCAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
Db 541 AIEVDGRVYHVLGDAQAQPRNLTKRHWPCAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
Qy 601 HRCYIILENDTVQCDDLDLYKSLQAWKHKLHIHDEIETLQNKIKNLRVGRGHLKKRPEEC 660
Db 601 HRCYIILENDTVQCDDLDLYKSLQAWKHKLHIHDEIETLQNKIKNLRVGRGHLKKRPEEC 660
Qy 661 DCHKI SYHTQHKGR LKRGSS LHP FRKGLQEKQVWLLREQKRLKLLKGLQNDTC 720

```

Db 661 DCHKISYHTQHKGLKRGSSLPFRKGLQEKDKVLLRQKQKKLKLKBLQNNDTC 720  
QY 721 SMPGLCTFHDNDHQTAPFTWLTGPFCACTSANNTYWCRTINETHNPLCFEATGELE 780  
Db 721 SMPGLCTFHDNDHQTAPFTWLTGPFCACTSANNTYWCRTINETHNPLCFEATGELE 780  
QY 781 YFDLNTDYPQLMNAVNTLDRDVLNQLRVQLMELRSCGYKQCNPRTRNMDLGLKDGGSYE 840  
Db 781 YFDLNTDYPQLMNAVNTLDRDVLNQLRVQLMELRSCGYKQCNPRTRNMDLGLKDGGSYE 840  
QY 841 QYQFQRKWPENKRPSSKSLGQLWEGWG 870  
Db 841 QYQFQRKWPENKRPSSKSLGQLWEGWG 870

## RESULT 8

US-09-989-722-84

; Sequence 84, Application US/09989722

; Patent No. US20020072067A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.

; APPLICANT: Baker, Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Deenoyers, Luc

; APPLICANT: Eston, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Fong, Sherman

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin L.

; APPLICANT: Kijavini, Ivar J.

; APPLICANT: Napier, Mary A.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; FILE REFERENCE: P2730P1063

; CURRENT APPLICATION NUMBER: US/09/989,722

; CURRENT FILING DATE: 2001-11-19

; PRIOR APPLICATION NUMBER: 60/049787

; PRIOR FILING DATE: 1997-06-16

; PRIOR APPLICATION NUMBER: 60/062250

; PRIOR FILING DATE: 1997-10-17

; PRIOR APPLICATION NUMBER: 60/065186

; PRIOR FILING DATE: 1997-11-12

; PRIOR APPLICATION NUMBER: 60/065311

; PRIOR FILING DATE: 1997-11-13

; PRIOR APPLICATION NUMBER: 60/066770

; PRIOR FILING DATE: 1997-11-24

; PRIOR APPLICATION NUMBER: 60/075945

; PRIOR FILING DATE: 1998-02-25

; PRIOR APPLICATION NUMBER: 60/078910

; PRIOR FILING DATE: 1998-03-20

; PRIOR APPLICATION NUMBER: 60/083322

; PRIOR FILING DATE: 1998-04-28

; PRIOR APPLICATION NUMBER: 60/084600

; PRIOR FILING DATE: 1998-05-07

; PRIOR APPLICATION NUMBER: 60/087106

; PRIOR FILING DATE: 1998-05-28

; PRIOR APPLICATION NUMBER: 60/087607

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/087609

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/089908

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088021  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088025  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088026  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088028  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088029  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088030  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088734  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088742  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088810  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088826  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088858  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088861  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088876  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/089105  
; PRIOR FILING DATE: 1998-06-12  
; PRIOR APPLICATION NUMBER: 60/089440  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089512  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089514  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089532  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089538  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089598  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089599  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089600  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089653  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089801  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089907  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089908

; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/089947  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/089948  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/089952  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/090246  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090252  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090254  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090349  
 ; PRIOR FILING DATE: 1998-06-23  
 ; PRIOR APPLICATION NUMBER: 60/090355  
 ; PRIOR FILING DATE: 1998-06-23  
 ; PRIOR APPLICATION NUMBER: 60/090429  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090431  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090435  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090444  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090445  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090472  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090535  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090540  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090542  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090557  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090676  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090678  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090690  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090694  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090695  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090696  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090862  
 ; PRIOR FILING DATE: 1998-06-26  
 ; PRIOR APPLICATION NUMBER: 60/090863  
 ; PRIOR FILING DATE: 1998-06-26  
 ; PRIOR APPLICATION NUMBER: 60/091360  
 ; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091478  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091544  
 ; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091519  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091626  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091633  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091978  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/091982  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/092192  
 ; PRIOR FILING DATE: 1998-07-09

Query Match 99.3%; Score 4719.5; DB 9; Length 867;

Best Local Similarity 99.5%; Pred. No. 0;  
 Matches 866; Conservative 0; Mismatches 1; Indels 3; Gaps 1;  
 Qy 1 MGPPSLVLCILSATVPSLLGGSSAFSLSHRLKGRQRRNRIRPNIIILVLTDDQDVEIGS 60  
 Db 1 MGPPSLVLCILSATVPSLLGGSSAFSLSHRLKGRQRRNRIRPNIIILVLTDDQDVEIGS 60  
 Qy 61 MQVNNKTRIMEQGGTHFINAFVTPMCCPSSSILTKYVHNHTYNNENCSPPSQA 120  
 Db 61 MQVNNKTRIMEQGGTHFINAFVTPMCCPSSSILTKYVHNHTYNNENCSPPSQA 120  
 Qy 121 QHESRTFAVILNSTGTRTAPFOKYLNEYNGSVPPGKWEVGLLKNSRFFNYTLCRNGVK 180  
 Db 121 QHESRTFAVILNSTGYRTAFFGKYLNEYNGSVPPGKWEVGLLKNSRFFNYTLCRNGVK 180  
 Qy 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRVLAVISHAAPHGPEDSAPQYSRLFP 240  
 Db 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRVLAVISHAAPHGPEDSAPQYSRLFP 240  
 Qy 241 NASOHITPSINYAPNDPKHMIMRYTGMKPIMHETNNMLQRKLOTILMSVDDSMETIYNN 300  
 Db 241 NASOHITPSINYAPNDPKHMIMRYTGMKPIMHETNNMLQRKLOTILMSVDDSMETIYNN 300  
 Qy 301 LVETGELDNITYIVYTADHGYHIGQFGLVKGSMRYEPDIRVPFYVGRGNVAGCLNPHIV 360  
 Db 301 LVETGELDNITYIVYTADHGYHIGQFGLVKGSMRYEPDIRVPFYVGRGNVAGCLNPHIV 360  
 Qy 361 LNIDLAPTILDIAGLDIPADMDGKSILKLDTERPVNRPHLKKQKRVWRDSFLVERGKLL 420  
 Db 361 LNIDLAPTILDIAGLDIPADMDGKSILKLDTERPVNRPHLKKQKRVWRDSFLVERGKLL 420  
 Qy 421 HKRDNDKVDQAQENFLPKYQVRVDLCORAEVCTACEOLGOKWOCVEDATGKLKHKCKGP 480  
 Db 421 HKRDNDKVDQAQENFLPKYQVRVDLCORAEVCTACEOLGOKWOCVEDATGKLKHKCKGP 480  
 Qy 481 MRLGGSRALSNLVPKYGGSEACTCDSDGYKLSLAGRRKKLFKKYKASYRSTRSV 540  
 Db 481 MRLGGSRALSNLVPKYGGSEACTCDSDGYKLSLAGRRKKLFKKYKASYRSTRSV 540  
 Qy 541 ALEVDRVYHVLGDAAPRNLTKRHWFGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600  
 Db 541 ALEVDRVYHVLGDAAPRNLTKRHWFGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600  
 Qy 601 HRCYILENDTVQCDLDLYKSLQAWKHLDHETLQNKIKNLRVGRHLKKRPEEC 660  
 Db 601 HRCYILENDTVQCDLDLYKSLQAWKHLDHETLQNKIKNLRVGRHLKKRPEEC 660  
 Qy 661 DCHKISYHTQHKGRKLRGSSLHPPFKGLQEKOKVYLLREOQKXKLRKLKRLQNNDT 720  
 Db 661 DCHKISYHTQHKGRKLRGSSLHPPFKGLQEKOKVYLLREOQKXKLRKLKRLQNNDT 720  
 Qy 721 SMPGLTCFTHDQHWOTAPFWTLGPFCACTSANNNYWCMTINETHNFLCFEATGFL 780  
 Db 721 SMPGLTCFTHDQHWOTAPFWTLGPFCACTSANNNYWCMTINETHNFLCFEATGFL 780  
 Qy 781 YFDLNTDYPQLMNAVNTLDRVNLQHLVQLMELRSCKYKQCNPRTRNMDLGLKDGSSYE 840  
 Db 781 YFDLNTDYPQLMNAVNTLDRVNLQHLVQLMELRSCKYKQCNPRTRNMDLGLKDGSSYE 840  
 Qy 841 QYRQFORRWKPEMKRPSKSLQGLWEGWEG 870  
 Db 838 QYRQFORRWKPEMKRPSKSLQGLWEGWEG 867

RESULT 9  
 US-09-989-723-84  
 ; Sequence 84, Application US/09989723  
 ; Patent No. US20020072092A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ashkenazi, Avi J.  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Botstein, David  
 ; APPLICANT: Desnoyers, Luc

APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerlitsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas P.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2730P1C62  
CURRENT APPLICATION NUMBER: US/09/989,723  
CURRENT FILING DATE: 2001-11-19  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444

;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090472  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090535  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090540  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090542  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090557  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090676  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090678  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090690  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090694  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090695  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090696  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090862  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091478  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091544  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091626  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091633  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091978  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/092182  
;; PRIOR FILING DATE: 1998-07-09

Query Match 99.3%; Score 4719.5; DB 9; Length 867;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 866; Conservative 0; Mismatches 1; Indels 3; Gaps 1;  
  
QY 1 MGPPSLVLCLSATVFSLLGSSAFSLSHRLKGRFQDRNRIRNIIILVLTDDQDVELGS 60  
DB 1 MGPPSLVLCLSATVFSLLGSSAFSLSHRLKGRFQDRNRIRNIIILVLTDDQDVELGS 60  
  
QY 61 MQWNNKTRIMEQGGTHINAFVITPMPCCPSRSSILTKGYVHNHTYTNNECSSPSQQA 120  
DB 61 MQWNNKTRIMEQGAHINAFVITPMPCCPSRSSILTKGYVHNHTYTNNECSSPSQQA 120  
  
QY 121 QHESFTFAYLNTGYRTAFPGKYLNENYNGSVYPPGKWKVGLLKNRFRNYITLCRNGVK 180  
DB 121 QHESFTFAYLNTGYRTAFPGKYLNENYNGSVYPPGKWKVGLLKNRFRNYITLCRNGVK 180  
  
QY 181 EKHGSDYKDYLTDLITDNDVSFRTSKMYPHRPLVMI SHAAPHGPDSPAPQYSLFP 240  
DB 181 EKHGSDYKDYLTDLITDNDVSFRTSKMYPHRPLVMI SHAAPHGPDSPAPQYSLFP 240  
  
QY 241 NASQHTPESYNAPNPKDGHINRYTGPMPKPIHMEFTNNLQKRLQTLMSVDDSMETIYNM 300  
DB 241 NASQHTPESYNAPNPKDGHINRYTGPMPKPIHMEFTNNLQKRLQTLMSVDDSMETIYNM 300  
  
QY 301 LVETGELONTYIVYTADHGYHIGQFLVKGKSMPEYEDIRVPFYVRGPNVEAGCLNPHIV 360

DB 301 LVETGELONTYIVYTADHGYHIGQFLVKGKSMPEYEDIRVPFYVRGPNVEAGCLNPHIV 360  
QY 361 LNIDLAFTILDIAGLDIPADMDGKSILKLLDTERPVRNRFHLKKQKRVWRDSFLVERGKLL 420  
DB 361 LNIDLAFTILDIAGLDIPADMDGKSILKLLDTERPVRNRFHLKKQKRVWRDSFLVERGKLL 420  
QY 421 HKRNDKYDAQSENFPLPKYQRYKOLCQRAEYQTACEQLGQKMQCQVEDATGKLKHLCKGKP 480  
DB 421 HKRNDKYDAQSENFPLPKYQRYKOLCQRAEYQTACEQLGQKMQCQVEDATGKLKHLCKGKP 480  
QY 481 MRLGGSRALSNLVPKYVGGSEACTCDSGDYKLSLAGRRKLLFKKYYKASYVRSRSRSV 540  
DB 481 MRLGGSRALSNLVPKYVGGSEACTCDSGDYKLSLAGRRKLLFKKYYKASYVRSRSRSV 540  
QY 541 AJEVDGRVYHVLGDAQAQPRNLTKRHWPAGPEDDQDKGGDPSTGGGLPDYSAANPIKVT 600  
DB 541 AJEVDGRVYHVLGDAQAQPRNLTKRHWPAGPEDDQDKGGDPSTGGGLPDYSAANPIKVT 600  
QY 601 HRCYILENDTVQCDLDLYKSLQANKDHKLHIDHETLONKIKNLREVRGHLKKRPREEC 660  
DB 601 HRCYILENDTVQCDLDLYKSLQANKDHKLHIDHETLONKIKNLREVRGHLKKRPREEC 660  
QY 661 DCHKISYHTQHKGRLLKRGSSLHPFRKGLQEKQKVKVLLREQKRRKLLKRLQNNDTG 720  
DB 661 DCHKISYHTQHKGRLLKRGSSLHPFRKGLQEKQKVKVLLREQKRRKLLKRLQNNDTG 720  
QY 721 SMPGLTCTHNDNQHWOTAPFTLGPFCACTSANNNTYCMETINETHNLFCEPATGFE 780  
DB 721 SMPGLTCTHNDNQHWOTAPFTLGPFCACTSANNNTYCMETINETHNLFCEPATGFE 780  
QY 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVMELRSCKGYKQCNPRTRNMDLGLKDGGSYE 840  
DB 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVMELRSCKGYKQCNPRTRNMDLGLKDGGSYE 837  
  
QY 841 QYRQFORRWKPEMKRPPSSKSLGQLWEGWEG 870  
DB 838 QYRQFORRWKPEMKRPPSSKSLGQLWEGWEG 867

RESULT 10  
US-09-989-299-84  
; Sequence 84, Application US/09989279  
; Patent No. US20020072496A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; TITLE OF INVENTION: Acids Encoding the Same  
; FILE REFERENCE: P2730PIC56  
; CURRENT APPLICATION NUMBER: US/09/989,279

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

[illegible]

1 PRIOR FILING DATE: 1998-06-02  
2 PRIOR APPLICATION NUMBER: 60/087759  
3 PRIOR FILING DATE: 1998-06-02  
4 PRIOR APPLICATION NUMBER: 60/087827  
5 PRIOR FILING DATE: 1998-06-03  
6 PRIOR APPLICATION NUMBER: 60/088021  
7 PRIOR FILING DATE: 1998-06-04  
8 PRIOR APPLICATION NUMBER: 60/088025  
9 PRIOR FILING DATE: 1998-06-04  
10 PRIOR APPLICATION NUMBER: 60/088026  
11 PRIOR FILING DATE: 1998-06-04  
12 PRIOR APPLICATION NUMBER: 60/088028  
13 PRIOR FILING DATE: 1998-06-04  
14 PRIOR APPLICATION NUMBER: 60/088029  
15 PRIOR FILING DATE: 1998-06-04  
16 PRIOR APPLICATION NUMBER: 60/088030  
17 PRIOR FILING DATE: 1998-06-04  
18 PRIOR APPLICATION NUMBER: 60/088033  
19 PRIOR FILING DATE: 1998-06-04  
20 PRIOR APPLICATION NUMBER: 60/088326  
21 PRIOR FILING DATE: 1998-06-04  
22 PRIOR APPLICATION NUMBER: 60/088167  
23 PRIOR FILING DATE: 1998-06-05  
24 PRIOR APPLICATION NUMBER: 60/088202  
25 PRIOR FILING DATE: 1998-06-05  
26 PRIOR APPLICATION NUMBER: 60/088212  
27 PRIOR FILING DATE: 1998-06-05  
28 PRIOR APPLICATION NUMBER: 60/088217  
29 PRIOR FILING DATE: 1998-06-05  
30 PRIOR APPLICATION NUMBER: 60/088655  
31 PRIOR FILING DATE: 1998-06-09  
32 PRIOR APPLICATION NUMBER: 60/088734  
33 PRIOR FILING DATE: 1998-06-10  
34 PRIOR APPLICATION NUMBER: 60/088738  
35 PRIOR FILING DATE: 1998-06-10  
36 PRIOR APPLICATION NUMBER: 60/088742  
37 PRIOR FILING DATE: 1998-06-10  
38 PRIOR APPLICATION NUMBER: 60/088810  
39 PRIOR FILING DATE: 1998-06-10  
40 PRIOR APPLICATION NUMBER: 60/088824  
41 PRIOR FILING DATE: 1998-06-10  
42 PRIOR APPLICATION NUMBER: 60/088826  
43 PRIOR FILING DATE: 1998-06-10  
44 PRIOR APPLICATION NUMBER: 60/088858  
45 PRIOR FILING DATE: 1998-06-11  
46 PRIOR APPLICATION NUMBER: 60/088861  
47 PRIOR FILING DATE: 1998-06-11  
48 PRIOR APPLICATION NUMBER: 60/088876  
49 PRIOR FILING DATE: 1998-06-11  
50 PRIOR APPLICATION NUMBER: 60/089105  
51 PRIOR FILING DATE: 1998-06-12  
52 PRIOR APPLICATION NUMBER: 60/089440  
53 PRIOR FILING DATE: 1998-06-16  
54 PRIOR APPLICATION NUMBER: 60/089512  
55 PRIOR FILING DATE: 1998-06-16  
56 PRIOR APPLICATION NUMBER: 60/089514  
57 PRIOR FILING DATE: 1998-06-16  
58 PRIOR APPLICATION NUMBER: 60/089532  
59 PRIOR FILING DATE: 1998-06-17  
60 PRIOR APPLICATION NUMBER: 60/089538  
61 PRIOR FILING DATE: 1998-06-17  
62 PRIOR APPLICATION NUMBER: 60/089598  
63 PRIOR FILING DATE: 1998-06-17  
64 PRIOR APPLICATION NUMBER: 60/089599  
65 PRIOR FILING DATE: 1998-06-17  
66 PRIOR APPLICATION NUMBER: 60/089600  
67 PRIOR FILING DATE: 1998-06-17  
68 PRIOR APPLICATION NUMBER: 60/089653  
69 PRIOR FILING DATE: 1998-06-17  
70 PRIOR APPLICATION NUMBER: 60/089801  
71 PRIOR FILING DATE: 1998-06-18  
72 PRIOR APPLICATION NUMBER: 60/089907  
73 PRIOR FILING DATE: 1998-06-18  
74 PRIOR APPLICATION NUMBER: 60/089908  
75 PRIOR FILING DATE: 1998-06-18  
76 PRIOR APPLICATION NUMBER: 60/089947  
77 PRIOR FILING DATE: 1998-06-19  
78 PRIOR APPLICATION NUMBER: 60/089948  
79 PRIOR FILING DATE: 1998-06-19  
80 PRIOR APPLICATION NUMBER: 60/089952  
81 PRIOR FILING DATE: 1998-06-19  
82 PRIOR APPLICATION NUMBER: 60/090246  
83 PRIOR FILING DATE: 1998-06-22  
84 PRIOR APPLICATION NUMBER: 60/090252  
85 PRIOR FILING DATE: 1998-06-22  
86 PRIOR APPLICATION NUMBER: 60/090254  
87 PRIOR FILING DATE: 1998-06-22  
88 PRIOR APPLICATION NUMBER: 60/090349  
89 PRIOR FILING DATE: 1998-06-23  
90 PRIOR APPLICATION NUMBER: 60/090355  
91 PRIOR FILING DATE: 1998-06-23  
92 PRIOR APPLICATION NUMBER: 60/090429  
93 PRIOR FILING DATE: 1998-06-24  
94 PRIOR APPLICATION NUMBER: 60/090431  
95 PRIOR FILING DATE: 1998-06-24  
96 PRIOR APPLICATION NUMBER: 60/090435  
97 PRIOR FILING DATE: 1998-06-24  
98 PRIOR APPLICATION NUMBER: 60/090444  
99 PRIOR FILING DATE: 1998-06-24  
100 PRIOR APPLICATION NUMBER: 60/090445  
101 PRIOR FILING DATE: 1998-06-24  
102 PRIOR APPLICATION NUMBER: 60/090472  
103 PRIOR FILING DATE: 1998-06-24  
104 PRIOR APPLICATION NUMBER: 60/090535  
105 PRIOR FILING DATE: 1998-06-24  
106 PRIOR APPLICATION NUMBER: 60/090540  
107 PRIOR FILING DATE: 1998-06-24  
108 PRIOR APPLICATION NUMBER: 60/090542  
109 PRIOR FILING DATE: 1998-06-24  
110 PRIOR APPLICATION NUMBER: 60/090557  
111 PRIOR FILING DATE: 1998-06-24  
112 PRIOR APPLICATION NUMBER: 60/090676  
113 PRIOR FILING DATE: 1998-06-25  
114 PRIOR APPLICATION NUMBER: 60/090678  
115 PRIOR FILING DATE: 1998-06-25  
116 PRIOR APPLICATION NUMBER: 60/090690  
117 PRIOR FILING DATE: 1998-06-25  
118 PRIOR APPLICATION NUMBER: 60/090694  
119 PRIOR FILING DATE: 1998-06-25  
120 PRIOR APPLICATION NUMBER: 60/090695  
121 PRIOR FILING DATE: 1998-06-25  
122 PRIOR APPLICATION NUMBER: 60/090696  
123 PRIOR FILING DATE: 1998-06-25  
124 PRIOR APPLICATION NUMBER: 60/090862  
125 PRIOR FILING DATE: 1998-06-26  
126 PRIOR APPLICATION NUMBER: 60/090863  
127 PRIOR FILING DATE: 1998-06-26  
128 PRIOR APPLICATION NUMBER: 60/091360  
129 PRIOR FILING DATE: 1998-07-01  
130 PRIOR APPLICATION NUMBER: 60/091478  
131 PRIOR FILING DATE: 1998-07-02  
132 PRIOR APPLICATION NUMBER: 60/091544  
133 PRIOR FILING DATE: 1998-07-01  
134 PRIOR APPLICATION NUMBER: 60/091519  
135 PRIOR FILING DATE: 1998-07-02  
136 PRIOR APPLICATION NUMBER: 60/091626  
137 PRIOR FILING DATE: 1998-07-02  
138 PRIOR APPLICATION NUMBER: 60/091633  
139 PRIOR FILING DATE: 1998-07-02  
140 PRIOR APPLICATION NUMBER: 60/091978  
141 PRIOR FILING DATE: 1998-07-07  
142 PRIOR APPLICATION NUMBER: 60/091982  
143 PRIOR FILING DATE: 1998-07-07  
144 PRIOR APPLICATION NUMBER: 60/092182  
145 PRIOR FILING DATE: 1998-07-09



```

Query Match      99.3%; Score 4719.5; DB 9; Length 867;
Beat Local Similarity 99.5%; Pred. No. 0;
Matches 866; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

Qy 1  MCPPSLVLCCLISATVFSLLGSSAFSLSHRLKGRFORDRNRINPIILVLTDDQDELGS 60
Db 1  MCPPSLVLCCLISATVFSLLGSSAFSLSHRLKGRFORDRNRINPIILVLTDDQDELGS 60

Qy 61  MQVMNKRIRIMEQGTHFINAFVITPMCCPSRSSILTKGYVHNHTYTNNENCSPPSQWA 120
Db 61  MQVMNKRIRIMEQGAHFINAFVITPMCCPSRSSILTKGYVHNHTYTNNENCSPPSQWA 120

Qy 121  QHESFTAVLNSTGYRTAFGKYLNEVNGSVVPGKWKGLKNSRFNYTLCRGVK 180
Db 121  QHESFTAVLNSTGYRTAFGKYLNEVNGSVVPGKWKGLKNSRFNYTLCRGVK 180

Qy 181  EXHGSYSDKYDLTDLITNDVSYSFFRTSKMYPHRPVLMVISHAAPHGPEDSAPOYSRLF 240
Db 181  EXHGSYSDKYDLTDLITNDVSYSFFRTSKMYPHRPVLMVISHAAPHGPEDSAPOYSRLF 240

Qy 241  NASQHITPSYNAPNPDGHWIMRYTGPMPKPIHMEFTNMLQKRLOTLSVDDSMETIYNN 300
Db 241  NASQHITPSYNAPNPDGHWIMRYTGPMPKPIHMEFTNMLQKRLOTLSVDDSMETIYNN 300

Qy 301  LVETGELDNTYIVYTADHGHIYHGFGLVKGKSMPEYFDIRVPFYVGRGNVEAGCLNPHIV 360
Db 301  LVETGELDNTYIVYTADHGHIYHGFGLVKGKSMPEYFDIRVPFYVGRGNVEAGCLNPHIV 360

Qy 361  LNIDLAFTILDLAGLDIPADMKGKSLKLLDTERPVNRFHLLKKXKRVWRDSFLVERGKLL 420
Db 361  LNIDLAFTILDLAGLDIPADMKGKSLKLLDTERPVNRFHLLKKXKRVWRDSFLVERGKLL 420

Qy 421  HKRDNDKVDQAEENFLPKYQVKOLCORABYOTACEQLGQKWCQVEDATGKLHKCKGP 480
Db 421  HKRDNDKVDQAEENFLPKYQVKOLCORABYOTACEQLGQKWCQVEDATGKLHKCKGP 480

Qy 481  MRLGGSRALSNLVPKYGGGSEACTCDSDGYKLSLAGRRKKLFKKYKASTVRSRSIRSV 540
Db 481  MRLGGSRALSNLVPKYGGGSEACTCDSDGYKLSLAGRRKKLFKKYKASTVRSRSIRSV 540

Qy 541  ATEVDGRVTHVGLGDAAPRNITKRWHPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
Db 541  ATEVDGRVTHVGLGDAAPRNITKRWHPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600

Qy 601  HRCYILENDTVQCDLDLYKSLQAKWQKHLIDHEIETLQNKILNREVRGHLKKRPEEC 660
Db 601  HRCYILENDTVQCDLDLYKSLQAKWQKHLIDHEIETLQNKILNREVRGHLKKRPEEC 660

Qy 661  DCHKISYHTQHKGRKLKRGSSLHPRFKGLQEKVDKQVLLREQRKKKLLKRLQNNDTC 720
Db 661  DCHKISYHTQHKGRKLKRGSSLHPRFKGLQEKVDKQVLLREQRKKKLLKRLQNNDTC 720

Qy 721  SMPGLTCFTHDNOHWOTAFPTWLTGPFCACTSANNTYMCMTTINETHNPLFCEPATGFE 780
Db 721  SMPGLTCFTHDNOHWOTAFPTWLTGPFCACTSANNTYMCMTTINETHNPLFCEPATGFE 780

Qy 781  YFDLNTDTPQLMNAVNTLDRVLNLQHLVQMLRELRSCKGYKQCNPRTRNMDLGLKDGGSYE 840
Db 781  YFDLNTDTPQLMNAVNTLDRVLNLQHLVQMLRELRSCKGYKQCNPRTRNMDLGLKDGGSYE 840

Qy 841  QYRQFORRWKPEMWRPSSKSLQGLWEGWG 870
Db 838  QYRQFORRWKPEMWRPSSKSLQGLWEGWG 867

```

RESULT 12

```

US-09-989-731-84
; Sequence 84, Application US/09989731
; Patent No. US20020103125A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David

```

```

; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC70
; CURRENT APPLICATION NUMBER: US/09/989,731
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202

```

```

; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24

```

```

; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09
; PRIOR FILING DATE: 1998-07-09

Query Match          99.3%; Score 4719.5; DB 10; Length 867;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 866; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

Qy/ 1 MGPPLVLCILSNVFSILGSSAFSLSHRLKGRFORDRNIRPNILVLTDDQDVELGS 60
Db 1 MGPPLVLCILSNVFSILGSSAFSLSHRLKGRFORDRNIRPNILVLTDDQDVELGS 60
Qy 61 MQVWNTKRIMEOGGTHFINAFVTTMCCPSRSSILTKYVHNHTYTNNECSPSQA 120
Db 61 MQVWNTKRIMEOGGTHFINAFVTTMCCPSRSSILTKYVHNHTYTNNECSPSQA 120
Qy 121 QHSRTFAVYLNSTGYRTAFFGKYLMNEVNGSVPPGKWEVGLLNKSRFNYTLCRNGVK 180
Db 121 QHSRTFAVYLNSTGYRTAFFGKYLMNEVNGSVPPGKWEVGLLNKSRFNYTLCRNGVK 180
Qy 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKKYPHRPVLWVISHAAPHGSDSAPQYSLFF 240
Db 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKKYPHRPVLWVISHAAPHGSDSAPQYSLFF 240
Qy 241 NASCHITPSYNYAPNPDKHWIMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNM 300
Db 241 NASCHITPSYNYAPNPDKHWIMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNM 300

```

QY 301 LVETGELONTYIVYTADHGYHIGQGLVKGKSNPYEFDIRPFFYVRGPNVEAGCLNPHIV 360  
Dd |||||  
QY 301 LVETGELONTYIVYTADHGYHIGQGLVKGKSNPYEFDIRPFFYVRGPNVEAGCLNPHIV 360  
Dd |||||  
QY 361 LNIIDLAPTILDIAGLDIPADMDGKSIILKLLDTERPVNRFHLKKGKRVWRDSDLVERGKLL 420  
Dd |||||  
QY 361 LNIIDLAPTILDIAGLDIPADMDGKSIILKLLDTERPVNRFHLKKGKRVWRDSDLVERGKLL 420  
Dd |||||  
QY 421 HKRDNKDVAEENFLPKYQRYKDLQRAEYQTAQCEQLGQKQWQVEDATGKLKHLKCKGP 480  
Dd |||||  
QY 421 HKRDNKDVAEENFLPKYQRYKDLQRAEYQTAQCEQLGQKQWQVEDATGKLKHLKCKGP 480  
Dd |||||  
QY 481 MRLGGSRALSNLVPKYGGGSEACTCDSDYKLSLAGRRKLLPKKKYKASVYRSRSIRSV 540  
Dd |||||  
QY 481 MRLGGSRALSNLVPKYGGGSEACTCDSDYKLSLAGRRKLLPKKKYKASVYRSRSIRSV 540  
Dd |||||  
QY 541 AIEVDGRVYHVLGDAQAQPNRLTKRHWPAGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600  
Dd |||||  
QY 541 AIEVDGRVYHVLGDAQAQPNRLTKRHWPAGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600  
Dd |||||  
QY 601 HRCYILENDTVQCDLDLYKSLQAWKHKLHIDHETITLQNKILREVRGHLKKGKPEEC 660  
Dd |||||  
QY 601 HRCYILENDTVQCDLDLYKSLQAWKHKLHIDHETITLQNKILREVRGHLKKGKPEEC 660  
Dd |||||  
QY 661 DCHKISYHTQHKGRLLKRGSSLHPRKGLQEKDKVWLLREQKRRKKLLKRLQNNDDTC 720  
Dd |||||  
QY 661 DCHKISYHTQHKGRLLKRGSSLHPRKGLQEKDKVWLLREQKRRKKLLKRLQNNDDTC 720  
Dd |||||  
QY 721 SMPGLTCFTHDNOHQWTAFTWLTGPFCACTSANNTYWCMTINETHNPLFCFATGPLE 780  
Dd |||||  
QY 721 SMPGLTCFTHDNOHQWTAFTWLTGPFCACTSANNTYWCMTINETHNPLFCFATGPLE 780  
Dd |||||  
QY 781 YFDLNTDTPQLMNAVNTLDRDVLNQLHVOLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840  
Dd |||||  
QY 781 YFDLNTDTPQLMNAVNTLDRDVLNQLHVOLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840  
Dd |||||  
QY 841 QYRQFORRWKPEMKRPPSSKSLQOLMEGWEG 870  
Dd |||||  
QY 838 QYRQFORRWKPEMKRPPSSKSLQOLMEGWEG 867  
Dd |||||

RESULT 13

US-09-989-732-84  
; Sequence 84, Application US/09989732  
; Patent No. US20020123463A1  
; GENERAL INFORMATION:  
; APPLICANT: Aehkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; ACIDS Encoding the Same  
; FILE REFERENCE: P2730PIC57

; CURRENT APPLICATION NUMBER: US/09/989,732  
; CURRENT FILING DATE: 2001-11-19  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/087106  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088021  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088025  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088026  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088028  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088029  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088030  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088734  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088742  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088810  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088826  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088858  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088861  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088876

; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/089105  
; PRIOR FILING DATE: 1998-06-12  
; PRIOR APPLICATION NUMBER: 60/089440  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089512  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089514  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089532  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089538  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089598  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089599  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089600  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089653  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089801  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089907  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089908  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089947  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/089948  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/089952  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/090246  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090252  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090254  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090349  
; PRIOR FILING DATE: 1998-06-23  
; PRIOR APPLICATION NUMBER: 60/090355  
; PRIOR FILING DATE: 1998-06-23  
; PRIOR APPLICATION NUMBER: 60/090429  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090431  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090435  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090444  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090445  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090472  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090535  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090540  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090542  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090557  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090676  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090678  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090690  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090694  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090695  
; PRIOR FILING DATE: 1998-06-25

; PRIOR APPLICATION NUMBER: 60/090696  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090862  
; PRIOR FILING DATE: 1998-06-26  
; PRIOR APPLICATION NUMBER: 60/090863  
; PRIOR FILING DATE: 1998-06-26  
; PRIOR APPLICATION NUMBER: 60/091360  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091478  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091544  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091519  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091626  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091633  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091978  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/091982  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/092182  
; PRIOR FILING DATE: 1998-07-09

Query Match 99.3%; Score 4719.5; DB 10; Length 867;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 866; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

Qy 1 MGPPSLVLCILLSATVPSILGSSAFSLHRLKGRFDRNRIRPNILVLTDQDVELGS 60  
Db 1 MGPPSLVLCILLSATVPSILGSSAFSLHRLKGRFDRNRIRPNILVLTDQDVELGS 60

Qy 61 MQVNMKTRIMEGGGTHFINAFVTTMCCPSRSSILTGKYVHNHTYNNENCSSPSWQA 120  
Db 61 MQVNMKTRIMEGGGAHFINAFVTTMCCPSRSSILTGKYVHNHTYNNENCSSPSWQA 120

Qy 121 QHESRTFAVYLNSTGYRTAFPGKYLNEYNGSVYPPGKEMVGLLKNRSFYNTLCRNGVK 180  
Db 121 QHESRTFAVYLNSTGYRTAFPGKYLNEYNGSVYPPGKEMVGLLKNRSFYNTLCRNGVK 180

Qy 181 EKHGSDYSKOYLTDLITNDOSVSFFRTSKQWYPHRPVLNVI SHAAHPGSDSAPQYSRLFP 240  
Db 181 EKHGSDYSKOYLTDLITNDOSVSFFRTSKQWYPHRPVLNVI SHAAHPGSDSAPQYSRLFP 240

Qy 241 NASQHITPSYNTAPNDKWIIMRYTGPMPKPIHMEFTNMLQRKRLQTLMSVDDSMETIYNM 300  
Db 241 NASQHITPSYNTAPNDKWIIMRYTGPMPKPIHMEFTNMLQRKRLQTLMSVDDSMETIYNM 300

Qy 301 LVETGELDNTYIVYTADHGVHIGQGLVKGSMPYEFDIRVPFYVRGPNVEAGCLNPHIV 360  
Db 301 LVETGELDNTYIVYTADHGVHIGQGLVKGSMPYEFDIRVPFYVRGPNVEAGCLNPHIV 360

Qy 361 LNTDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPFLHKKQVWRDSEFLVERGKLL 420  
Db 361 LNTDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPFLHKKQVWRDSEFLVERGKLL 420

Qy 421 HKRDNDKVDQAQENFLPKYQRYKDLICORAEYOTACEOLGQKWCQVEDATGKLLHKCKGP 480  
Db 421 HKRDNDKVDQAQENFLPKYQRYKDLICORAEYOTACEOLGQKWCQVEDATGKLLHKCKGP 480

Qy 481 MRLGSSRALSNLVPKYTGQSGSEACTCDSGDYKLSLAGRRKKLFKKYKASYVRSRSIRSV 540  
Db 481 MRLGSSRALSNLVPKYTGQSGSEACTCDSGDYKLSLAGRRKKLFKKYKASYVRSRSIRSV 540

Qy 541 AIEVDGRVTHVGLGDAAPRNLTKRHWPCAPEDQDDKGGDFSGTGGGLPDYSAAPIKVT 600  
Db 541 AIEVDGRVTHVGLGDAAPRNLTKRHWPCAPEDQDDKGGDFSGTGGGLPDYSAAPIKVT 600

Qy 601 HRCYILENDTVOCDLLYKSLQAKDHLHI DHEIETLQNKIKNLRVGRHGLKKRPREC 660  
Db 601 HRCYILENDTVOCDLLYKSLQAKDHLHI DHEIETLQNKIKNLRVGRHGLKKRPREC 660

661	Qy	DCHKISYHTQHKGRUKHGRSSLIHPFKGLQEKQKVWLLRQKQKKKLRKULLKGLQNNDC	720
661	Db	DCHKISYHTQHKGRUKHGRSSLIHPFKGLQEKQKVWLLRQKQKKKLRKULLKGLQNNDC	720
721	Qy	SMPGLTCTFHDNOHWOTAPFMTLGPCCACTSANNTYWCMTINETHNPFCBFAFGLE	780
721	Db	SMPGLTCTFHDNOHWOTAPFMTLGPCCACTSANNTYWCMTINETHNPFCBFAFGLE	780
781	Qy	YFDLNTDPYQLMNAVNTLDRDVLNQLVQLMELRSCGKYQCNCPRTRNMDLGLKDGGSYE	840
781	Db	YFDLNTDPYQLMNAVNTLDRDVLNQLVQLMELRSCGKYQCNCPRTRNMDL----DGGSYE	837
841	Qy	QYRQFORRKWPEMKPSSKSLGQLWEGWEG	870
838	Db	QYRQFORRKWPEMKPSSKSLGQLWEGWEG	867

## RESULT 14

```

US-09-991-073-84
; Sequence 84, Application US/09991073
; Patent No. US2002012756A1
GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C15
CURRENT APPLICATION NUMBER: US/09/991,073
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02

```

; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/089908  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/089947  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/089948  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/089952  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/090246  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090252  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090254  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090349  
 ; PRIOR FILING DATE: 1998-06-23  
 ; PRIOR APPLICATION NUMBER: 60/090355  
 ; PRIOR FILING DATE: 1998-06-23  
 ; PRIOR APPLICATION NUMBER: 60/090429  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090431  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090435  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090444  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090445  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090472  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090535  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090540  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090542  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090557  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090676  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090678  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090690  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090694  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090695  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090696  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090862  
 ; PRIOR FILING DATE: 1998-06-26  
 ; PRIOR APPLICATION NUMBER: 60/090863  
 ; PRIOR FILING DATE: 1998-06-26  
 ; PRIOR APPLICATION NUMBER: 60/091360  
 ; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091478  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091544  
 ; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091519  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091626  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091633  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091978  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/091982  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/092182  
 ; PRIOR FILING DATE: 1998-07-09

Query Match 99.3%; Score 4719.5; DB 10; Length 867;  
 Best Local Similarity 99.5%; Pred. No. 0;  
 Matches 866; Conservative 0; Mismatches 1; Indels 3; Gaps 1;  
  
 Qy 1 MGPPSLVLCILLSATVFSLLGGSSAFUSHLRLKGRFORDRNRIRPNILVLTDODVELGS 60  
 Db 1 MGPPSLVLCILLSATVFSLLGGSSAFUSHLRLKGRFORDRNRIRPNILVLTDODVELGS 60  
 Qy 61 MQVMNKTRIMEOGGTHFINAFVTTMCCPSRSSILTKYVHNHNTYTNNECSSPSWQA 120  
 Db 61 MQVMNKTRIMEOGGAHFINAFVTTMCCPSRSSILTKYVHNHNTYTNNECSSPSWQA 120  
 Qy 121 QHESRTFAVYLNSTGYRTAFPGKYLNEYNGSVYPPGKKEWVGLLKNRSFYNTLCEGVK 180  
 Db 121 QHESRTFAVYLNSTGYRTAFPGKYLNEYNGSVYPPGKKEWVGLLKNRSFYNTLCEGVK 180  
 Qy 181 EKHSYSDYKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAPHGSDSAPQYSLFP 240  
 Db 181 EKHSYSDYKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAPHGSDSAPQYSLFP 240  
 Qy 241 NASQHITPSYNYAPNPKHIMRYTGMKPIHMEFTNMLQKGLQTLMSVDDSMETIYNN 300  
 Db 241 NASQHITPSYNYAPNPKHIMRYTGMKPIHMEFTNMLQKGLQTLMSVDDSMETIYNN 300  
 Qy 301 LVETGELDNITYIYTTADHGYHICQFGLVKGKSNPYEFDIRVPFYVGRPNVAGCLNPHIV 360  
 Db 301 LVETGELDNITYIYTTADHGYHICQFGLVKGKSNPYEFDIRVPFYVGRPNVAGCLNPHIV 360  
 Qy 361 LNIDLAPTILDIAGLDIPADMDOKSILKLLDTERPVNRFHLKKKMRVWRDSEFLVERGKLL 420  
 Db 361 LNIDLAPTILDIAGLDIPADMDOKSILKLLDTERPVNRFHLKKKMRVWRDSEFLVERGKLL 420  
 Qy 421 HKRDNDKVDQAQENFLPKYQVXKDLQRAEYQACBQLGOKWQCVEDATGKLLHKCKGP 480  
 Db 421 HKRDNDKVDQAQENFLPKYQVXKDLQRAEYQACBQLGOKWQCVEDATGKLLHKCKGP 480  
 Qy 481 MRIGGSRALSNLVPKYVGGSEACTCDSDYKLSLAGRRKLLPKKCYKASYVRSRSIRSV 540  
 Db 481 MRIGGSRALSNLVPKYVGGSEACTCDSDYKLSLAGRRKLLPKKCYKASYVRSRSIRSV 540  
 Qy 541 AIEVDGRVYHVGLGDAQAQPRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600  
 Db 541 AIEVDGRVYHVGLGDAQAQPRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600  
 Qy 601 HRCYILENDTVQCDLDLYKSLQAWKDHKLHIDHEIETLQNKIKNLRVGRHLKKGRPEEC 660  
 Db 601 HRCYILENDTVQCDLDLYKSLQAWKDHKLHIDHEIETLQNKIKNLRVGRHLKKGRPEEC 660  
 Qy 661 DCHKISYHTQHKGRLLKRGSSLLHPRFKGLQEKDKVLLREOKKKKLLKLLKRLQNNDT 720  
 Db 661 DCHKISYHTQHKGRLLKRGSSLLHPRFKGLQEKDKVLLREOKKKKLLKLLKRLQNNDT 720  
 Qy 721 SMPGLTCFTHDNQHWOTAPFWTLGPPFCACTSANNTYWCMTINETHNLFCEPATGFLE 780  
 Db 721 SMPGLTCFTHDNQHWOTAPFWTLGPPFCACTSANNTYWCMTINETHNLFCEPATGFLE 780  
 Qy 781 YFDLNTDPPYOLMNAVNTLDRDVLNQLHVQLMELRSCKYKQCNPRTNMDLGLKQGSYE 840  
 Db 781 YFDLNTDPPYOLMNAVNTLDRDVLNQLHVQLMELRSCKYKQCNPRTNMDLGLKQGSYE 840  
 Qy 841 QYRQFQRKWPMPKRPSSKSLGQWEGWEG 870  
 Db 838 QYRQFQRKWPMPKRPSSKSLGQWEGWEG 867

RESULT 15  
 US-09-990-442-84  
 ; Sequence 84, Application US/09990442  
 ; Patent No. US20020132252A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ashkenazi, Avi J.  
 ; APPLICANT: Baker, Kevin P.

APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Fertara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Geritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730PIC8  
CURRENT APPLICATION NUMBER: US/09/990,442  
CURRENT FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088126  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435

1	PRIOR FILING DATE: 1998-06-24	
2	PRIOR APPLICATION NUMBER: 60/090444	
3	PRIOR FILING DATE: 1998-06-24	
4	PRIOR APPLICATION NUMBER: 60/090445	
5	PRIOR FILING DATE: 1998-06-24	
6	PRIOR APPLICATION NUMBER: 60/090472	
7	PRIOR FILING DATE: 1998-06-24	
8	PRIOR APPLICATION NUMBER: 60/090535	
9	PRIOR FILING DATE: 1998-06-24	
10	PRIOR APPLICATION NUMBER: 60/090540	
11	PRIOR FILING DATE: 1998-06-24	
12	PRIOR APPLICATION NUMBER: 60/090542	
13	PRIOR FILING DATE: 1998-06-24	
14	PRIOR APPLICATION NUMBER: 60/090557	
15	PRIOR FILING DATE: 1998-06-24	
16	PRIOR APPLICATION NUMBER: 60/090676	
17	PRIOR FILING DATE: 1998-06-25	
18	PRIOR APPLICATION NUMBER: 60/090678	
19	PRIOR FILING DATE: 1998-06-25	
20	PRIOR APPLICATION NUMBER: 60/090690	
21	PRIOR FILING DATE: 1998-06-25	
22	PRIOR APPLICATION NUMBER: 60/090694	
23	PRIOR FILING DATE: 1998-06-25	
24	PRIOR APPLICATION NUMBER: 60/090695	
25	PRIOR FILING DATE: 1998-06-25	
26	PRIOR APPLICATION NUMBER: 60/090696	
27	PRIOR FILING DATE: 1998-06-25	
28	PRIOR APPLICATION NUMBER: 60/090862	
29	PRIOR FILING DATE: 1998-06-26	
30	PRIOR APPLICATION NUMBER: 60/090863	
31	PRIOR FILING DATE: 1998-06-26	
32	PRIOR APPLICATION NUMBER: 60/091360	
33	PRIOR FILING DATE: 1998-07-01	
34	PRIOR APPLICATION NUMBER: 60/091478	
35	PRIOR FILING DATE: 1998-07-02	
36	PRIOR APPLICATION NUMBER: 60/091544	
37	PRIOR FILING DATE: 1998-07-01	
38	PRIOR APPLICATION NUMBER: 60/091519	
39	PRIOR FILING DATE: 1998-07-02	
40	PRIOR APPLICATION NUMBER: 60/091626	
41	PRIOR FILING DATE: 1998-07-02	
42	PRIOR APPLICATION NUMBER: 60/091633	
43	PRIOR FILING DATE: 1998-07-02	
44	PRIOR APPLICATION NUMBER: 60/091978	
45	PRIOR FILING DATE: 1998-07-07	
46	PRIOR APPLICATION NUMBER: 60/091982	
47	PRIOR FILING DATE: 1998-07-07	
48	PRIOR APPLICATION NUMBER: 60/092182	
49	PRIOR FILING DATE: 1998-07-09	

Query Match	99.3%	Score 4719.5	DB 10	Length 867
Best Local Similarity	99.5%	Pred. No. 0		
Matches 866	Conservative 0	Mismatches 1	Indels 3	Gaps 1
QY	1	MGPPSLVLCLLSATVFSLLGGSSAFLSHRLKGRFORDRENIRPNILVLTTDDQDVELGS	60	
Db	1	MGPPSLVLCLLSATVFSLLGGSSAFLSHRLKGRFORDRENIRPNILVLTTDDQDVELGS	60	
QY	61	MOVNMKTRIMEOGSTHEFINAVFTTPCCSRSSILTGKVVHNNHTYTNNECSPSSWQA	120	
Db	61	MOVNMKTRIMEOGGAHFINAVFTTPCCSRSSILTGKVVHNNHTYTNNECSPSSWQA	120	
QY	121	QHESTRFAVYLNSTGYRTAFTEGKLYNEYSYVPPGKWEVGLLKNSRFPNYTLCRNGVK	180	
Db	121	QHESTRFAVYLNSTGYRTAFTEGKLYNEYSYVPPGKWEVGLLKNSRFPNYTLCRNGVK	180	
QY	181	EXHGSQDYSKDYLTDLITNDSVSFRTSKQMPVHRPVLWVLSHAAPGHPEDSAPOYSRLFP	240	
Db	181	EXHGSQDYSKDYLTDLITNDSVSFRTSKQMPVHRPVLWVLSHAAPGHPEDSAPOYSRLFP	240	
QY	241	NASQHITPSYNYAPNPDGHWIMRYTGPMPKPIHMEFTNMLQRKQLTILMSYDDSMETIYNM	300	
Db	241	NASQHITPSYNYAPNPDGHWIMRYTGPMPKPIHMEFTNMLQRKQLTILMSYDDSMETIYNM	300	

Qy	301	LVETGELDNTIIVYTA	DGHYHIGOFGLVKGSKMPYBFDIRVPYVGRGPNVVEAGCLNPHIV	360
Db	301	LVETGELDNTIIVYTA	DGHYHIGOFGLVKGSKMPYBFDIRVPYVGRGPNVVEAGCLNPHIV	360
Qy	361	LNI DLAPTILDIAGL	DPADMDGKSILKLJDTERPYNRPHLKKKRWDRDSFLVBERGKLL	420
Db	361	LNI DLAPTILDIAGL	DPADMDGKSILKLJDTERPYNRPHLKKKRWDRDSFLVBERGKLL	420
Qy	421	HKRDNDKYDAQEEN	FLPKYQRVKDLCORASYQTACEQLGQKWQCVEDATGKLLKHKCKGP	480
Db	421	HKRDNDKYDAQEEN	FLPKYQRVKDLCORAEYQTACEQLGQKWQCVEDATGKLLKHKCKGP	480
Qy	481	MLGGSRALSNLVPKY	QCGSEACTCDSGDYKLSLAGRRKLLFKKKYKASYVRSRSIRSV	540
Db	481	MLGGSRALSNLVPKY	QCGSEACTCDSGDYKLSLAGRRKLLFKKKYKASYVRSRSIRSV	540
Qy	541	ALEVDRVYHVGLG	DAQAQPNRLTKRHWFGAPBDQDDKDGDFSGTGGLPDYSAANPIKVT	600
Db	541	ALEVDRVYHVGLG	DAQAQPNRLTKRHWFGAPBDQDDKDGDFSGTGGLPDYSAANPIKVT	600
Qy	601	HRCYILENDTYQC	DLLYKSLQAWKDKHLHDIETLQNKIKNLEVRGHLKKKGRPEEC	660
Db	601	HRCYILENDTYQC	DLLYKSLQAWKDKHLHDIETLQNKIKNLEVRGHLKKKGRPEEC	660
Qy	661	DCHKISYHTQHGRL	KHRGSSLIHPRFGLOEKDKWLLREQKRKKLKLKLQNNDDTC	720
Db	661	DCHKISYHTQHGRL	KHRGSSLIHPRFGLOEKDKWLLREQKRKKLKLKLQNNDDTC	720
Qy	721	SMPLGTCFTHDN	QHWQAPFWTGLGPFCACTSANNTYWCWRTINETHNLFLPCBFATGFLE	780
Db	721	SMPLGTCFTHDN	QHWQAPFWTGLGPFCACTSANNTYWCWRTINETHNLFLPCBFATGFLE	780
Qy	781	YFDLNDTPYQLMNA	VNTLDRDVLNQLHVLMELRSCKYKQCNCPRTRNMDLGLKGGGSYE	840
Db	781	YFDLNDTPYQLMNA	VNTLDRDVLNQLHVLMELRSCKYKQCNCPRTRNMDLGLKGGGSYE	840
Qy	841	QYRQFORRWKPMK	RPPSSKSLGULWEGWG	870
Db	838	QYRQFORRWKPMK	RPPSSKSLGULWEGWG	867

Search completed: October 15, 2003, 13:19:02  
Job time : 83 secs





```

Db      448 QAKCYTVEDRWRYKCR-----DNWSDQSC-----R 476
QY      520 KILFKKKYKASVRSRSIRSAIEVDGRVYHVLGDAAQPRNLTKRWHPGAPEDQDDXG 579
Db      477 KK-----REI-----SNYDDDDI 489
QY      580 GDFSGTGLPDYSAANPIKVTWRCYILENDTVQCDLDLYKSLQAMKDKHLHIDHIEITLQ 639
Db      490 DEF-----LTYADRENFSGEHMYQGEFEDSGEV-----G 519
QY      640 NKIKNLRVRGHLKKRPECCDKHISYHTQKGLKRGSLHFRKGLQEKQKVMLLR 699
Db      520 EELDGRSKRGILSK-----CSCSR-----NVSHPIK-----LLE 549
QY      700 EQRKRLKRLKRLKRLQND-----TCSMGLCTGTHDNQHWOTAPFW-----TLGPPFCACTSAN 753
Db      550 QRMKKHLYKRYKKPQNGSLKPKDCSLPQMNCKFTHTASHMKTPPLWPELGEFCFCQNCN 609
QY      754 NNTYCMRTINETHNLFCEPATGFEYFDLNTDPVQLMNAVNTLDRVLNQLHVQLMEL 813
Db      610 NNTYCLRTKNTNHLFYCFEFTFISFYDNTDPQLINAVYSLDIGVLSQLRLNL 669
QY      814 RSCGYKQC 822
Db      670 RKCKN-RQC 677

RESULT 2
KUHUGU
N-Acetylglucosamine-6-sulfatase (EC 3.1.6.14) precursor - human
N:Alternate names: Chondroitinsulfatase, N-acetyl-D-glucosamine-6-sulfate 6-sulfohydrolase
C:Species: Homo sapiens (man)
C:Date: 22-Nov-1993 #sequence_revision 27-Oct-1995 #text_change 24-Sep-1999
C:Accession: S27164; A31672
R:Robertson, D.A.; Freeman, C.; Morris, C.P.; Hopwood, J.J.
Biochem. J. 288, 539-544, 1992
A:Title: A cDNA clone for human glucosamine-6-sulphatase reveals differences between aryl
A:Reference number: S27164; MUID:93098807; PMID:1463457
A:Accession: S27164
A:Molecule type: mRNA
A:Residues: 1-552 <ROB>
A:Cross-references: GB:Z12173; EMBL:M23657; NID:G31866; PIDN:CAA78164.1; PID:G31867
A>Note: parts of this sequence, including the amino end of the mature protein, were deter
R:Robertson, D.A.; Freeman, C.; Nelson, P.V.; Morris, C.P.; Hopwood, J.J.
Biochem. Biophys. Res. Commun. 157, 218-224, 1988
A:Title: Human glucosamine-6-sulfatase cDNA reveals homology with steroid sulfatase.
A:Reference number: A31672; MUID:89061714; PMID:3196333
A:Accession: A31672
A:Molecule type: mRNA
A:Residues: 178-552 <RO2>
A:Cross-references: GB:Z12173; EMBL:M23657; NID:G31866
C:Genetics:
A:Gene: GDB:GNS
A:Cross-references: GDB:120006; OMIM:252940
A:Map position: 12q14-12q14
A>Note: Defects in this gene can cause mucopolysaccharidosis type III D, Sanfilippo D di
C:Function:
A:Description: hydrolyzes N-acetyl-D-glucosamine 6-sulfate units in heparan sulfate and
C:Superfamily: animal sulfatase
C:Keywords: glycoprotein; lysosomal storage disease; lysosome; Sanfilippo disease; sulfu
F:1-43/Domain: signal sequence #status predicted <SIG>
F:44-552/Product: N-acetylglucosamine-6-sulfatase #status predicted <MAT>
F:91/Modified site: 3-oxoalanine (Cys) #status predicted
F:111,117,183,198,210,279,317,362,387,405,449,480/Binding site: carbohydrate (Asn) (cova
F:422/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match      18.1%; Score 858.5; DB 1; Length 552;
Best Local Similarity 27.2%; Pred. No. 4,4e-54;
Matches 256; Conservative 76; Mismatches 174; Indels 355; Gaps 18;

QY      4 PSILVCLLSAT--VPSLLGSSAFSLSHRLKGRFORDRNRIRPNILVLTDDQDVELGSM 61

```

```

Db      24 PALLLVGLGCLGVFGVGAAGTR-----RPNVLLLTDDQDEVLGGM 64
QY      62 QVNMKTRRIMEQGTTHFINAFVTTMPCPSRSRSLTGRKYVHHNTYTN--NENCSFSPSQ 119
Db      65 TPLKTKKALIGEMGTFSSAYVPSALCCPSRASILTGRKYPHNHVVNNTLEGNCSSKSNQ 124
QY      120 AQHESHTPAVYLS--TCYRTAPFGKYLNEYG-----SYVPGWKWVGLLKNRPFNY 172
Db      125 KIQEPNTFAILLFSCMGYQTFAGKYLYNEYGAPDAGGLEHVPGLGWSYVALEKNSKYNY 184
QY      173 TLCNRNGKCKHSGDYSDKYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAPHGEDSA 232
Db      185 TLSINGKARKGKGENYSDYLTDLVLANVSLDFLDYKSNF---EPFFMMIATPAPHSPTAA 241
QY      233 PQVSRLLPNNASQHTPSYNYAPNDKHWIMRYT-GPMKPIHMEFTNMLQKRLQTLMSVD 291
Db      242 PQYQKAFQNVFAPRNKNFN-IHGTNKHMLIROAKTPTMTNSSIQFLDNAPRKRWQTLISVD 300
QY      292 DSMETIYNMLVETGELDNNTYIVYTDHGYHIGOPGLVKCKSMPEYEDIRVPFYVRGPVNE 351
Db      301 DLVEKLVKRLKLEFTEGLNNTYIFYTSNGYHTGQFSLPIDKRLQYEPDIKVPILLVRPGIK 360
QY      352 AGCLNPHIVNLIDLAFTILDIAGLDI-PADMDGKSLKLLDTERPVNRPFLKKQKRVWRD 410
Db      361 PNQTSKMLVANIDLGPTILDIAGYDLNKQMDGMSLLPIL---RGASNL-----TWRS 410
QY      411 SFLVERKLLHKKRDNDKVDQAQENFLPKYQVKNLCORAEYQVACQOLQKQKQVEDATG 470
Db      411 DVLVE----- 415
QY      471 KKLHKCKGPMRLGGSRALSNLVKYVGSGSEACTCDSDGYKLSLAGRRKKLKKYKAS 530
Db      416 -----YQSGS----- 420
QY      531 YVRSRSIRSAIEVDGRVYHVLGDAAQPRNLTKRWHPGAPEDQDDKOGDFTSGTGLPD 590
Db      421 -----RNV----- 424
QY      591 YSAANPIKVTWRCYILENDTVQCDLDLYKSLQAMKDKHLHIDHIEITLQNKIKNLRVGR 650
Db      425 ----- 424
QY      651 HLKKRPECCDKHISYHTQKGLKRGSLHFRKGLQEKQKVMLLREQKRRKKLRL 710
Db      425 -----DPTCPSL----- 432
QY      711 LKRLQNNDTCSMPGLT-CPTHNDQHWOTAPFWTLGPFCACTSANNTYVQMTINETHNF 769
Db      433 -----PGVSQCF-----PDCVCEAYNNTYACVRTMSALMNL 464
QY      770 LFCEP--ATGFEYLPDLNTPVQLMNAVNTLDRVLNQLHVQLMELRSCGK 818
Db      465 QYCFDDQEVFVEVYNTADPDQITNIATIDPELLGQWYRLMMLQSCSG 515

RESULT 3
T46577
aryl sulfatase (EC 3.1.6.1) [validated] - Neurospora crassa
C:Species: Neurospora crassa
C:Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 04-Mar-2000
C:Accession: T46577
R:Baker, D.L.; Paletta, J.V.
submitted to the EMBL Data Library, February 1997
A:Description: Molecular characterization of the aryl sulfatase gene of Neurospora cras
A:Reference number: Z23090
A:Accession: T46577
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-639 <BAK>
A:Cross-references: EMBL:U89492; PIDN:AAC02716.1
A:Experimental source: strain wild type 74-OR23-1A
C:Genetics:

```

A:Description: [EC 3.1.6.1] [validated, MUID:94222055]; hydrolyzes sulfate esters in r  
A:Note: required for mineralization of sulfate  
A:Note: remarkably insensitive towards detergents like dodecyl sulfate  
C:Superfamily: plant sulfatase  
C:Keywords: glycoprotein; periplasmic space; sulfuric ester hydrolase  
F1-/2/Domain: signal sequence #status predicted <Sig>  
F23-/649/Product: arylsulfatase #status experimental <MAT>  
F72-/Modified site: 3-oxoalanine (Cys) #status experimental

Query Match	9.5%;	Score	452.5;	DB	2;	Length	649;
Best Local Similarity	27.4%;	Pred. No.	1.3e-24;				
Matches	149;	Conservative	82;	Mismatches	170;	Indels	143;
Gaps	24;						

  

QY	6	LVLCLLSATVSLGGSSAFLSHRLKGRFQDRRRIRPNIIILVLTDDODDELGSM-QV	63
DB	7	VALCLLG---PAALTAANA---HQ-----RPNFVVFITDDQDGTNSTHPY	47
QY	64	MNKTRRIMBOGGTHFINAFVTTMCCPSRSIIITGKYVHNHNNTYTNNCSSP-----SW	118
DB	48	QPKLHEHIRYPCIBLKNVFVTTTPCCPSRTNLWRGQFSHN---TNFTDVLGPHGGYAKM	103
QY	119	QAQHESTRP-AVYLNSTGYRTAFGKYLNEY-----NGSYVPPGKEWVGLLKNSRF-YNTT	173
DB	104	KSLGIDKSYLPVNLQNLGYNTYYVVKELVDYSVSNYQNVPAGTDDALVPTPTFDYNNP	163
QY	174	LCRNGVKEKHGSYSDXYLTDLTNDVSFFRTSKOMYRHRVLMVISHAAPHGSDSA	232
DB	164	GFSRNGATP---NIYPGYSTDVLDKVAQIKTA---VAAGKDPYAIQISFIAPH-----	212
QY	233	POYSLRFPNASHITPSTYNTAPNP-DKHWTMRVYGPM-----KPIHMEF	275
DB	213	-TSTQIYFDPVANATKTFPPPIAPRPHWELFSDATLPECTSHKNLYEADVSDKPAWIRA	271
QY	276	TNMLQ-----BKRLQTLMSVDDSMETIYNNLVTGELDNTYIVYTDHGHVHIGQ	324
DB	272	LPLAQONRTYLEEYRLRLKRLASVDELDRVATLQEAGVLNDTYLYISADNGHVST	331
QY	325	FLGVKGSMPYEFDIRVPFYVRGPNVEAG-----CLNPHVILNIDLAPTILDIAG-----	374
DB	332	HRFGAGKVTADEDLRVFELIRGPIRASHSDKPAANSKVGLHVDFAPTILTLAGAGDQVG	391
QY	375	-----LDIPLADMCGKSLKLLDTERPYNRPHLKKRVRMRDSFLVE-----RCKLIH	421
DB	392	DKALDGTPLGUYANDCGN---LLADYPRPANH-----RNQPGFWMGWSDEVILH	438
QY	422	KRDNDKVDQAENFLPKYQ-----RVKDLQCORAEYQTACEQLGQKQWCVEDATGKLK	474
DB	439	-----HIPRYTNNSWKAVRVYD-----EDNQAWKLIVSCTNEREL	474
QY	475	HKCK	478
DB	475	YDLK	478

RESULT 5  
KJFM4  
arylsulfatase (EC 3.1.6.1) precursor - Chlamydomonas reinhardtii  
N/Alternate names: sulfatase  
C/Species: Chlamydomonas reinhardtii  
C/Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 24-Sep-1999  
C/Accession: JQ0310  
R:de Hostos, E.B.; Schilling, J.; Grossman, A.R.  
Mol. Gen. Genet. 218, 229-239, 1989  
A/Title: Structure and expression of the gene encoding the periplasmic arylsulfatase O  
A/Reference number: JQ0310; MUID:89384447; PMID:2476654  
A/Accession: JQ0310  
A/Molecule type: mRNA  
A/Residues: 1-646 <DEH>  
A/Cross-references: GB:X16180; NID:g18118; PIDN:CAA34302.1; PID:g18119; GB:X52304; NID  
A/Experimental source: strain cw15mt+  
A/Note: part of this sequence, including the amino end of the mature protein, was deter  
C/Comment: This enzyme is commonly produced by soil microorganisms and plays an import  
C/Superfamily: plant sulfatase

Matches 105; Conservative 72; Mismatches 154; Indels 81; Gaps 21;

Db 223

## Match

Gakawara, N.; Iasunaga, A.; Amura, S.; Shima, I.; Hattori, M.; Saitogawa, H.  
 DNA Res. 8, 11-22, 2001  
 A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic islands  
 A:Reference number: A99629; PMID:21156231; PMID:11258796  
 A:Accession: G90891  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-571 -RAY>  
 A:Cross-references: GB:BA000007; PIDN:BA035526.1; PID:g13361569; GSPDB:GN00154  
 A:Experimental source: strain O157:H7, substrain RIMD 0509952  
 C:Genetics:

C;Superfamily: animal sulfatase

[illegible]

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

五

[illegible]







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2003, 13:01:00 ; Search time 15 Seconds

(without alignments)  
2727.549 Million cell updates/sec

Title: US-09-970-287-2

Perfect score: 4751

Sequence: 1 MGPPSLVLCLSATVPSLLG.....PEMKRPSSKLGOLWEGWG 870

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	858.5	18.1	552	1 GL6S_HUMAN	P15586 homo sapien
2	856	18.0	559	1 GL6S_CAPIH	P50426 capra hircu
3	452.5	9.5	649	1 ARS_VOLCA	Q10723 volvox cart
4	405	8.5	647	1 ARS_CHLRE	P14217 chlamydomon
5	302.5	6.4	497	1 YIDJ_ECOLI	P31447 escherichia
6	256	5.4	512	1 BETC_RHIME	O69787 rhizobium m
7	251	5.3	560	1 YDEN_ECOLI	P77318 escherichia
8	238.5	5.0	583	1 STS_HUMAN	P08842 homo sapien
9	238	5.0	567	1 ARS_STRPU	P50473 strongyloce
10	234.5	4.9	550	1 IDS_HUMAN	P22304 homo sapien
11	232	4.9	551	1 ARS_HEMPU	P14000 hemitentrot
12	226	4.8	563	1 IDS_MOUSE	Q08890 mus musculu
13	218.5	4.6	522	1 GA6S_HUMAN	P34059 homo sapien
14	210.5	4.4	506	1 ARSB_MOUSE	P50428 mus musculu
15	210.5	4.4	533	1 ARSB_HUMAN	P15848 homo sapien
16	210.5	4.4	577	1 STS_RAT	P15589 rattus norv
17	203	4.3	535	1 ARSB_FELCA	P33727 felis silve
18	201.5	4.2	464	1 ARS_KLEAE	P20713 klebsiella
19	200	4.2	593	1 ARSD_HUMAN	P51689 homo sapien
20	198.5	4.2	591	1 ARSF_HUMAN	P54793 homo sapien
21	194.5	4.1	507	1 ARSA_HUMAN	P15289 homo sapien
22	189.5	4.0	624	1 STS_MOUSE	P50427 mus musculu
23	183.5	3.9	589	1 ARSE_HUMAN	P51690 homo sapien
24	176.5	3.7	502	1 SPHM_HUMAN	P51688 homo sapien
25	176	3.7	551	1 ASLA_ECOLI	P25549 escherichia
26	170	3.6	473	1 ARSB_RAT	P50430 rattus norv
27	155.5	3.3	535	1 ARS_PSEAE	P51691 pseudomonas
28	138.5	2.9	510	1 GPML_LISIN	Q92812 listeria in
29	137.5	2.9	510	1 GPML_LISNO	O8Y414 listeria mo
30	135.5	2.9	510	1 GPML_ANTSP	Q06464 antithamio
31	134	2.8	533	1 GPML_ANASP	Q8Y212 anabaena sp
32	127	2.7	285	1 ARSB_MOUSE	P50429 mus musculu
33	127	2.7	532	1 GPML_SYNY3	P74507 synechocyst

34	125.5	2.6	724	1	ATIL_VACCV	P24759 vaccinia vi
35	124	2.6	510	1	GPML_BACHD	Q9K716 bacillus ha
36	124	2.6	512	1	GPML_OCEIH	P59174 oceanobacil
37	123	2.6	638	1	YQGS_BACSU	P54496 bacillus su
38	123	2.6	2339	1	RPCI_PLAFA	P27625 plasmodium
39	122	2.6	510	1	GPML_CLOAB	Q97153 clostridium
40	122	2.6	534	1	GPML_PORPU	P51379 porphyra pu
41	118.5	2.5	875	1	DD10_HUMAN	Q13206 homo sapien
42	118	2.5	510	1	GPML_BACST	Q9X519 bacillus st
43	117	2.5	512	1	GPML_CLOPE	Q8XK22 clostridium
44	112.5	2.4	1435	1	EBAL_PLAFC	P19214 plasmodium
45	112	2.4	1616	1	VITI_CAEL	P55155 caenorhabdi

## ALIGNMENTS

RESULT 1				
ID	GL6S_HUMAN	STANDARD;	PRT;	552 AA.
AC	P15586;			
DT	01-APR-1990 (Rel. 14, Created)			
DT	01-APR-1993 (Rel. 25, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	N-acetylglucosamine-6-sulfatase precursor (EC 3.1.6.14) (G6S)			
DE	(Glucosamine-6-sulfatase).			
GN	GNS.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Endothelial cells;			
RX	MEDLINE=93098807; PubMed=1463457;			
RA	Robertson D.A., Freeman C., Morris C.P., Hopwood J.J.;			
RT	"A cDNA clone for human glucosamine-6-sulphatase reveals differences			
RT	between arylsulphatases and non-arylsulphatases.";			
RL	Biochem. J. 288:539-544 (1992).			
RN	(2)			
RP	SEQUENCE OF 178-552 FROM N.A., AND PARTIAL SEQUENCE.			
RX	MEDLINE=89061714; PubMed=3196333;			
RA	Robertson D.A., Freeman C., Nelson P.V., Morris C.P., Hopwood J.J.;			
RT	"Human glucosamine-6-sulfatase cDNA reveals homology with steroid			
RT	sulfatase.";			
RL	Biochem. Biophys. Res. Commun. 157:218-224 (1988).			
CC	!- CATALYTIC ACTIVITY: Hydrolysis of the 6-sulfate group of the N-			
CC	acetyl-D-glucosamine 6-sulfate units of heparan sulfate and			
CC	keratan sulfate.			
CC	!- SUBCELLULAR LOCATION: Lysosomal.			
CC	!- DISEASE: DEFECTS IN GNS RESULT IN THE ACCUMULATION OF PARTIALLY			
CC	DEGRADED HEPARAN SULFATE IN LYOSOMES CAUSING ORGANELLE CELL AN			
CC	TISSUE DISTORTION, ULTIMATELY LEADING TO THE LYOSOMAL STORAGE			
CC	DISORDER, MUCOPOLYSACCHARIDOSIS TYPE IIID (MPS-IIID) (ALSO KNOWN			
CC	AS SANFILIPPO D SYNDROME).			
CC	!- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collabor			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstat			
CC	the European Bioinformatics Institute. There are no restrictions on			
CC	use by non-profit institutions as long as its content is in no			
CC	modified and this statement is not removed. Usage by and for comm			
CC	entities requires a license agreement (See <a href="http://www.iesb-sib.ch/ann">http://www.iesb-sib.ch/ann</a>			
CC	or send an email to <a href="mailto:license@iesb-sib.ch">license@iesb-sib.ch</a> ).			
CC	EMBL; Z12173; CAA78164.1; -.			
DR	PIR; S27164; KJHUGU.			
DR	Genew; HGNC:4422; GNS.			
DR	MIM; 252940; -.			
DR	GO; GO:0008449; F:N-acetylglucosamine-6-sulfatase activity; TAS.			
DR	GO; GO:0006027; F:glycosaminoglycan catabolism; TAS.			
DR	InterPro; IPR000917; Sulfatase.			
DR	Pfam; PF00884; Sulfatase; 1.			



```

FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 487 487 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 559 AA; 62711 MW; 786CCDC48334A58 CRC64;

Query Match
Best Local Similarity 27.5%; Score 856; DB 1; Length 559;
Matches 231; Conservative 77; Mismatches 177; Indels 354; Gaps 20;

QY 3 PPSLVLCILSATVSLGGSSAFSLSHLKGKRFDRRNPNIILVLTDDQVELGSMQ 62
Db 29 PPPLLLLL-----LGGCLG-----VSGAAGSR---RPNVVLVLAQQDEVLGGM 72
QY 63 VMKTRIMEQGGTHFINAFVTPMCCPSRSSIITGKYVHNHTYTN--NENGSSSSQWA 120
Db 73 PLKTKKALIGEMGMFTFSAYVPSALCCPSRASILTGKYPHNHVNVNTLEGNCSSKQWK 132
QY 121 QHESRTFVYLNS--TGRTAFPGKYLEYNG-----SYPPGKWEVGLLKNSRFNYT 173
Db 133 IQEPNTPFPAILRSMCGYQTFAGKYLEYNGAPDAGGLGHVPLGWSYWYALEKSKYNT 192
QY 174 LCRNGVKEKGDYSKDYLTDLTNDVSF--FRTSKMYPHRPVLMVISHAAPHGPEDS 231
Db 193 LSIKGARKGENTSVDTLTVLANSLDFDYKNSN-----PPFMISTPAPHSPMTA 247
QY 232 APOYSRLFPNASQHITPSYNAFNPDKHWMRYT--GPMKPIHMEFTNMLQKRIQLTMSV 290
Db 248 APOYONAFQNVFAPRNKFN--IHGTNKHLLIRQAQKTPMNSSIQFLDNAPRERQWILL 306
QY 291 DDSNETYNMLVEGELDNVYIVTADGHHIGQFGLVKGSKMPEYEDIRVYVQGNV 350
Db 307 DDLVEKLVKRLFPNGELNNTYIFTSNGVHTGQFSLPIDKQQLYEDFDKVPVLVRGPI 366
QY 351 EAGCLNPHVILNIDLAPLIDAGLDI--PADMDCKSLKLLDTERPVNRPHLKQKQVVR 409
Db 367 KPNQTSKWLVAIDLGTILDIAGYGLNKTKQMGNSFLPII--RGASNL-----TWR 416
QY 410 DSFLVERGKLLHKRDNKVDQAEENFLPKYQVXKDCQRAEYOTACBQLQKQWQVEDAI 469
Db 417 SDVLVE-----RNVF-----422
QY 470 GKLKHLKCKGPMRLGGSEALSNLVPKYVGQSGEACTCDSDGYKLSLAGRKKLFKKKYKA 529
Db 423 -----YQEG-----427
QY 530 SYVRSRGSIRVAIEVDGRVYHVLGLDAAQPRNLKRWHPGAPEDQDKGDFSGTGGLP 589
Db 428 -----RNVF-----431
QY 590 DYSAANPIKVTHRCYILENTVQCDDLQYKSLQAWKQHLHIDHEIETLQNKIKULREVR 649
Db 432 -----431
QY 650 GHLKKRPEECCHKISVYTHQKGLKRGSSLLHPRKGLQKDKWLLREKQKKKLRK 709
Db 432 -----DPTCPSLS-----439
QY 710 LLKGLQNDTCSMPGLT--CFTHDQHWQTAFTWTLGPFCACTSANNTYWCMTINETHN 768
Db 440 -----PQVSOCP-----PDCVCEADYNNYACVVTMSLWN 470
QY 769 FLFCFEF--ATGFELEYFDLNTDPYQLMNAVNTLDRDVNLQHLVQMLERSCYKQCNPR 825
Db 471 LOYCEFDQOEVEVYVNLTAHPHOLNIAKSIDPELLGKNNYRLMLMQSCSGTCRTPR 529

RESULT 3
ARS_VOLCA
ID - ARS VOLCA STANDARD; PRT; 649 AA.
AC Q10723;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Arylsulfatase precursor (EC 3.1.1.6.1) (Aryl-sulfate sulphohydrolase).

```

```

OS Volvox carteri.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Volvocaceae; Volvox.
OX NCBI_TaxID=3067;
RN [1]
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.
RX STRAIN=f. Nagariensis / HK10;
RX MEDLINE=94222055; PubMed=8168504;
RA Hallmann A., Sumper M.;
RT "An inducible arylsulfatase of Volvox carteri with properties
RT suitable for a reporter-gene system. Purification, characterization
RT and molecular cloning.";
RL Eur. J. Biochem. 221:143-150(1994).
RN [2]
RP SEQUENCE OF 64-76, AND POST-TRANSLATIONAL MODIFICATION OF CYS-72.
RX MEDLINE=96283826; PubMed=8681943;
RA Selmer T., Hallmann A., Schmidt B., Sumper M., von Figura K.;
RT "The evolutionary conservation of a novel protein modification, the
RT conversion of cysteine to serinesemialdehyde in arylsulfatase from
RT Volvox carteri.";
RL Eur. J. Biochem. 238:341-345(1996).
CC -!- FUNCTION: Is commonly produced by soil microorganisms and plays an
CC important role in the mineralization of sulfates.
CC -!- CATALYTIC ACTIVITY: A phenol sulfate + H(2)O = a phenol + sulfate.
CC -!- ENZYME REGULATION: INHIBITED BY NA(3)BO(3) AND KCN. NO INHIBITION
CC BY SODIUM DODECYL SULFATE, EVEN AT HIGH CONCENTRATION.
CC -!- SUBCELLULAR LOCATION: Periplasmic.
CC -!- INDUCTION: By sulfur deprivation.
CC -!- MISCELLANEOUS: The enzyme is thermostable, exhibiting a
CC temperature optimum at 60 degrees Celsius. Its optimal pH is 8.0.
CC -!- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X77214; CAA54426.1; -.
DR PIR; S43229; S43229.
DR InterPro; IPR000917; Sulfatase.
DR Pfam; PF00884; Sulfatase; 1.
DR PROSITE; PS00523; SULFATASE_1; 1.
DR PROSITE; PS00149; SULFATASE_2; 1.
KW Hydrolase; Signal; Glycoprotein; Periplasmic.
FT SIGNAL 1 22
FT CHAIN 23 649 ARYLSULFATASE.
FT MOD_RES 72 72 2-AMINO-3-OXOPROPIONIC ACID.
FT DOMAIN 571 582 POLY-ALA.
FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 224 224 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 279 279 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 489 489 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 531 531 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 649 AA; 72287 MW; 0C23BFD7C43F7B9 CRC64;

Query Match
Best Local Similarity 27.4%; Score 452.5; DB 1; Length 649;
Matches 149; Conservative 82; Mismatches 170; Indels 143; Gaps 24;

QY 6 LVLCILSATVSLGGSSAFSLSHLKGKRFDRRNPNIILVLTDDQVELGSM--QV 63
Db 7 VALCLLG---FAALTAAAA---HQ-----RPNFVWIFTDDQDGIQNSTHPRY 47
QY 64 MNKTRIMEQGGTHFINAFVTPMCCPSRSSIITGKYVHNHTYTNENSCSSP-----SW 118
Db 48 QPKLHEHRYPGIELKNYFVTPVCCPSRTNLMRGQFSHN-----TNFDVLGPHGQYAKW 103
QY 119 QAQHSRTF-AVYLNSTGYRTAFPGKYLEY---NGSYVPPGKWEVGLLKNSRF--YNYT 173

```

Db	104	KS LGIDKSYLPWMLQNLQNTYYVYKGLVYSVSNYQNVPAGWTDIDALVYTFYTDYNNP	163
Qy	174	-LCRNGVKEGSDYSKDYLDTLINDSVSGFTTSKKYPHRPVLWVISHAAPHGPEDSA	232
Db	164	GFSRNGATP---NITYPGYSYTDVLDADKVAQIKTA--VAAGKEFYAQISPLAPH-----	212
Qy	233	POYSRLFPNASOHITPSNYAENP-DKHWMRYTGPW-----KPIHMEF	275
Db	213	-TSTQIYFDPVANAKTTFYPPIAPRHWELFSDATLPEGTSKNNLYEADVSDKPAWIRA	271
Qy	276	TNMLQ-----RKRLQTLMSVSDSMETIYNMLVETGELDNTIYVYVYADHGHGHCQ	324
Db	272	LPLAQONRTYLEEYVRLRSLASVDLDRVVATLQEAQVLNDTYLIYSADNGYHVGT	331
Qy	325	FGLVGKSKMPEEDIRVPYVYRGDNVEAG---CLNPHIVLNDLAPTILDIAG-----	374
Db	332	HRFGAGKVTAYDEDURDLPFLIRPGIRASHSDKSPANSKVLGVDFVDPATILTLAGAGDQVG	391
Qy	375	-----LDIDADMKGKSI LKLALDTRPPVNRPHLKKGRVWRDSDLVE-----RGKLLH	421
Db	392	DKALDGTPLGLYANDGN---LLADYPRPANH-----RNQOQGEFWGWSDEVHL	438
Qy	422	KRDNDKVDAAQENFLPKYQ-----RYKOLCORAEYQTACEQUGKQWQCVEDATGKML	474
Db	439	-----HIPRYTNNSWKAIVYD-----EDNQQAOKMLVYSTCTNEREL	474
Qy	475	HKCK	478
Db	475	YDLK	478

## RESULT 4

```

ARS_CHLRE
ARS_CHLRE
ID ARS_CHLRE STANDARD; PRT; 647 AA.
AC P14217;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Arylsulfatase precursor (EC 3.1.6.1) (Aryl-sulfate sulphohydrolase).
GN AS.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 22-55.
RC STRAIN=GW15;
RX MEDLINE=89384447; PubMed=2476654;
RA de Hostos E.L., Schilling J., Grossman A.R.;
RT *Structure and expression of the gene encoding the periplasmic
RT arylsulfatase of Chlamydomonas reinhardtii.";
RL Mol. Gen. Genet. 218:229-239(1989).
RN [2]
RP IDENTIFICATION OF PROBABLE FRAMESHIFTS.
RA Balroch A.;
RA Unpublished observations (AUG-1996).
CC -1- CATALYTIC ACTIVITY: A phenol sulfate + H(2)O = a phenol + sulfate.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- INDUCTION: By sulfur deprivation.
CC -1- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
CC -1- CAUTION: THIS IS A CONCEPTUAL TRANSLATION, FRAMESHIFTS HAD TO BE
CC INTRODUCED TO MAXIMIZE THE SIMILARITY WITH OTHER SULFATASES.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X16180; CAA34302.1; ALT FRAME.

```

[illegible]

## RESULT 5

## RESULTS

1033-PTA

ID	YIDJ
----	------

AC P3144

DT 01-JU

DT 01-JU

DT 16-OC

DE Putat

DE  
GN  
FUCAL  
YIND

GN  
OC

50 50 50

OC Bacte

OC Enter

OX NCBI\_

RN [1]

E



```

Db 158 --ITWQMEYDDVAFANOKUQLSRENDDESRRPWCULTVSTHHPDHYVARRKFWDLYE 215
Qy 241 NASQHTSYNAP-----NPKDHLMTYTGPMKPIHMEFTNM-----LQKR--LQTL 287
Db 216 DC-EHLTEVGAIPLEQPHSQRM-----LSCDYQFDTVEENRSTRAYFANI 266
Qy 288 MSVDDSMETINMLVETBELDNTYIVYADHGYHGGQGLVKGKSMPEFDIRVFFYVRG 347
Db 267 SYLDEKVGELIDTLRTRLMDLTLLFCSDHGMGLGERGL-WFKWNPFGSGARVPLMIAG 325
Qy 348 PNVEAGC-LNPHIVLNIDLAPTILDIAGL--DIPADMDGKSILKILD 391
Db 326 FGIAGLHLTP--TSNLDVPTLADLAGISLEVRPWTGVSVPWVN 371

RESULT 7
ID YDEN ECOLI STANDARD; PRT; 560 AA.
AC P77318; P78159;
DT 16-OCT-2001 (Rel. 40, Created)
DE 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative sulfatase yden precursor (EC 3.1.6.-).
GN YDEN OR B1498.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R.; Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA Kasai H., Kashimoto K., Kimura S., Kitagawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
RA Sampei G., Seki Y., Sivasubram S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
CC -!- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF000247; AAC74571.1; ALT INIT.
DR EMBL; D90791; BAA15169.1; ALT INIT.
DR EMBL; D90792; BAA15172.1; ALT_INIT.
DR EcoGene; EG13796; yden.
DR InterPro; IPR000917; Sulfatase.
DR Pfam; PF00884; Sulfatase; 1.
DR PROSITE; PS00523; SULFATASE_1; FALSE_NEG.
DR PROSITE; PS00149; SULFATASE_2; 1.
KW Hypothetical protein; Hydrolase; Signal; Complete proteome.
FT SIGNAL 1 29 POTENTIAL.
FT CHAIN 30 560 PUTATIVE SULFATASE YDEN.
FT ACT_SITE 185 185 POTENTIAL.

```

```

SQ SEQUENCE 560 AA; 62802 MW; 67ED7FF7696C7A9F CRC64;
Query Match 5.3%; Score 251; DB 1; Length 560;
Best Local Similarity 22.7%; Pred. No. 1.6e-10;
Matches 129; Conservative 75; Mismatches 187; Indels 178; Gaps 25;

Qy 18 LLGSGSAFUSHRLKGRFORDRNI-----RPNILVLTDQ----- 54
Db 18 LASGMAFAAHAADDVKLKATKTNVAFSDFTPTYSTKGKNNIIVLTMDDLGCGQLPDK 77
Qy 55 -----DVELGSMQVWNNKTRIMEOGGTHFINAFVTPMCCPERS 93
Db 78 GGFDPKTMENREVDTYKIGIDKATEAAQKSTPTLLSLMDEGVRTNGYVAHGSGPSRA 137
Qy 94 SILTGKYVHNHTYTNENCCSPSWQAQHE---SRTFAVIL-NSTGYRTAFPGK-YLNEY 148
Db 138 AIMGTRAPARFGVYSNTD-----AQDGIPLTFTFLPQLNHHGYTAAVKGWHLSKI 189
Qy 149 NGSYPPGKWEVGLLKNSRFY--NYTL--CRNGVKERHGSY----- 187
Db 190 SNVPVPED-----KQTRDYHDFNFTTFSABEQPNQRFDFMGFHAAGTAYYNSPSL 241
Qy 188 -----SKDYLTLITNDSVSFFRTSKMYPHPRLVLMVISHAAPHGPEDS-AP-QYSR 237
Db 242 FKXREVPKAGYISDQLTDEAIGVDRAKTL--DQPFMLYLAYNAPHLPNDNPADQYQK 299
Qy 238 LFPNASOHITPSYNYAPNPDGKIMRYTGPMPKPIHMEFTMLQRKRLQTLMSVDDSMETI 297
Db 300 QFNTGSQ--TADNYA-----SVYSVQCGVKRI 325
Qy 298 YNMLVETGELDNTYIVYADHGYH--GQF---GLVKG-KSMPEYEDIRVPFVVRGPNVEA 352
Db 326 LSQLKXNGYDNTIILFTSDNGAVTDGFLPLNGAKGYKSYQTPGTHPMPMMW----K 381
Qy 353 GCINP-----HIVLNIDLAPTILDIAGLDIPAD--MDGKSILKLLDTER---PVNRFHLKK 403
Db 382 GKLPQNYDKLISAMDFTALDAADISIPKDLKLDGVSLLPWLQDKKQGEHPHKLWIT 441
Qy 404 KMRVRDSTFLVERGKLLHK-----RDND--KYDAQENFL 436
Db 442 SYSHMFDEENIPFDNYHKFVRHOSDDYPHNPTEDLSQFSYTVRNNDSLVYTVNNQL 501
Qy 437 PKYQVWDLCCORAEYCTACEOLGQKQWCV 465
Db 502 GUY-KLTDLQKDNLAANPQVVKEXGV 529

RESULT 8
STS HUMAN
ID STS HUMAN STANDARD; PRT; 583 AA.
AC P08842;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Steryl-sulfatase precursor (EC 3.1.6.2) (Steroid sulfatase) (Steryl-
DE sulfate sulfohydrolase) (Arylsulfatase C) (ASC).
GN STS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89340479; PubMed=2668275;
RA Stein C., Hille A., Seidel J., Rijnbout S., Waheed A., Schmidt B.,
RA Geuze H., von Figura K.;
RT "Cloning and expression of human steroid-sulfatase. Membrane
RT topology, glycosylation, and subcellular distribution in BHK-21
RT cells.";
RL J. Biol. Chem. 264:13865-13872(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87187642; PubMed=3032454;

```

RA Yen P.H., Allen E., Marsh B., Mohandas T., Wang N., Taggart R.T.,  
RA Shapiro L.J.;  
RT "Cloning and expression of steroid sulfatase cDNA and the frequent  
RT occurrence of deletions in STS deficiency: implications for X-Y  
RT interchange.";  
RL Cell 49:443-454(1987).  
RN [3]  
RP SEQUENCE OF 134-174 AND 461-583 FROM N.A.  
RX MEDLINE=89077541; PubMed=3203382;  
RA Yen P.H., Marsh B., Allen E., Tsai S.P., Ellison J., Connolly L.,  
RA Neiswanger K., Shapiro L.J.;  
RT "The human X-linked steroid sulfatase gene and a Y-encoded  
RT pseudogene: evidence for an inversion of the Y chromosome during  
RT primate evolution.";  
RL Cell 55:1123-1135(1988).  
RN [4]  
RP SEQUENCE OF 22-45.  
RC TIGSUE=Liver;  
RX MEDLINE=89352671; PubMed=2765556;  
RA Kawano J.-I., Kofani T., Ohtaki S., Minamino N., Matsuo H., Oinuma T.,  
RA Aikawa E.;  
RT "Characterization of rat and human steroid sulfatases.";  
RL Biochim. Biophys. Acta 997:199-205(1989).  
RN [5]  
RP VARIANTS XLI LEU-341; ARG-372 AND TYR-446.  
RX MEDLINE=92170784; PubMed=1539590;  
RA Basler E., Grompe M., Parenti G., Yates J., Ballabio A.;  
RT "Identification of point mutations in the steroid sulfatase gene of  
RT three patients with X-linked ichthyosis.";  
RL Am. J. Hum. Genet. 50:483-491(1992).  
RN [6]  
RP VARIANTS XLI LEU-341; ARG-372; SER-372; ARG-444 AND TYR-446.  
RX MEDLINE=97400563; PubMed=9252398;  
RA Alperin E.S., Shapiro L.J.;  
RT "Characterization of point mutations in patients with X-linked  
RT ichthyosis. Effects on the structure and function of the steroid  
RT sulfatase protein.";  
RL J. Biol. Chem. 272:20756-20763(1997).  
RN [7]  
RP VARIANT XLI PRO-560.  
RX MEDLINE=20146224; PubMed=10679952;  
RA Sugawara T., Shimizu H., Hoshi N., Fujimoto Y., Nakajima A.,  
RA Fujimoto S.;  
RT "PCR diagnosis of X-linked ichthyosis: identification of a novel  
RT mutation (B560P) of the steroid sulfatase gene";  
RL Hum. Mutat. 15:296-296(2000).  
RN [8]  
RP VARIANT XLI ARG-380.  
RX MEDLINE=20304877; PubMed=10844566;  
RA Oyama N., Satoh M., Iwatsuki K., Kaneko F.;  
RT "Novel point mutations in the steroid sulfatase gene in patients with  
RT X-linked ichthyosis: transfection analysis using the mutated genes.";  
RL J. Invest. Dermatol. 114:1195-1199(2000).  
CC -!- FUNCTION: CONVERSION OF SULFATED STEROID PRECURSORS TO ESTROGENS  
CC DURING PREGNANCY.  
CC -!- CATALYTIC ACTIVITY: 3-beta-hydroxyandrost-5-en-17-one 3-sulfate +  
CC H(2O) = 3-beta-hydroxyandrost-5-en-17-one + sulfate.  
CC -!- SUBUNIT: Homodimer.  
CC -!- SUBCELLULAR LOCATION: MICROSMAL MEMBRANE. THE SEQUENCE SHOWS  
CC SEVERAL MEMBRANE-SPANNING DOMAINS THAT COULD SERVE TO ANCHOR THE  
CC PROTEIN IN THE MICROSMAL MEMBRANE.  
CC -!- DISEASE: Defects in STS are the cause of x-linked ichthyosis  
CC (XLI), a diskarization disorder characterized by the presence  
CC of prominent scales, especially on the neck, extremities, trunk,  
CC and buttocks. It affects approximately 1 of 2000-6000 males.  
CC -!- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC -----  
CC EMBL; J04964; AAA60597.1; -;  
DR EMBL; M16505; AAA60596.1; -;  
DR EMBL; M23945; AAA60598.1; -;  
DR EMBL; M23556; AAA60599.1; -;  
DR PIR; A32641; KJHUAC.  
DR HSSP; F15846; IFSU.  
DR Genew; HGNC:11425; STS.  
DR MIM; 308100; -;  
DR GO; GO:0005783; C:endoplasmic reticulum; TAS.  
DR GO; GO:0005768; C:endosome; TAS.  
DR GO; GO:0005794; C:Golgi apparatus; TAS.  
DR GO; GO:0005764; C:lysosome; TAS.  
DR GO; GO:0005624; C:membrane fraction; TAS.  
DR GO; GO:0005792; C:microsome; TAS.  
DR GO; GO:0005886; C:plasma membrane; TAS.  
DR GO; GO:0004773; F:steryl-sulfatase activity; TAS.  
DR GO; GO:0006544; P:epidermal differentiation; TAS.  
DR GO; GO:0006706; P:steroid catabolism; TAS.  
DR InterPro; IPR000917; Sulfatase.  
DR Pfam; PF00884; Sulfatase; 1.  
DR PROSITE; PS00523; SULFATASE 1; 1.  
DR PROSITE; PS00149; SULFATASE 2; 1.  
DR KW Hydrolase; Microsome; Transmembrane; Glycoprotein; Steroid metabolism;  
KW Pregnancy; Signal; Disease mutation.  
FT SIGNAL 1 21  
FT CHAIN 22 583  
FT MOD\_RES 75 75  
FT STERYL-SULFATASE.  
FT 2-AMINO-3-OKOPROPIONIC ACID (BY  
FT SIMILARITY).  
FT POTENTIAL.  
FT POTENTIAL.  
FT POTENTIAL.  
FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT S -> L (IN XLI; LOSS OF ACTIVITY).  
FT /FTID=VAR 007240.  
FT W -> R (IN XLI; LOSS OF ACTIVITY).  
FT /FTID=VAR 007241  
FT W -> S (IN XLI; LOSS OF ACTIVITY).  
FT /FTID=VAR 014020.  
FT G -> R (IN XLI).  
FT /FTID=VAR 014021.  
FT H -> R (IN XLI; LOSS OF ACTIVITY).  
FT /FTID=VAR 014022.  
FT C -> Y (IN XLI; LOSS OF ACTIVITY).  
FT /FTID=VAR 007242.  
FT Q -> P (IN XLI).  
FT /FTID=VAR 014023.  
FT A -> E (IN REF. 2).  
FT SEQUENCE 583 AA; 65492 MW; 74746AFA9D21A0A6 CRC64;  
SQ  
Query Match 5.0%; Score 238.5; DB 1; Length 583;  
Best Local Similarity 20.7%; Pred. No. 1.3e-09;  
Matches 126; Conservative 67; Mismatches 180; Indels 237; Gaps 23;  
QY 43 RPNILVLTDDDDVGLSGMVN-NXTR1-----MEOGTHFINAFVTPMCCPSRSIL 96  
DB 26 RPNILVLTDD--DLGIDPGCYGNTIITPNIDRLASGVKLTQHLAASPLCTPSRAAFM 83  
QY 97 TGYVHNHTYNNENCSPSWQ-----AQHESRTFAYLVNLTGYRTAFEGK 143  
DB 84 TGRY-----PVRSGWMSWRTGVFLFTASSGGLPTBITPAKLLKDGYSTALIGK 134  
QY 144 Y-----LNEY-----NGSYVPPGWKE----- 159  
DB 135 WHLGMSCHSKTDFFCHPLHGFNYFYGISLTNLRDKPGEGSVFTTGKRLVFLPLQIVG 194  
QY 160 -----WVGLKXSRFYNYTLCRNGVKRKHG 184  
DB 195 VTLLTLAALNCGLLHVPLGVFPFSLFLAALITLFLGLHFLYRPLNCFMRN--YBIIQ 252

```

QY 185 SDYSKDYLTDLTNDVSFRTSKMYPHRPVLVMSHAAPHGPDSPAPQYSLRFPNASQ 244
DB 253 QPMVDNLTLQRLTVEAQQIQRNTE-----TPFLVLVSLVHV-----TALFSSKDFAGKSQ 304
QY 245 HTTSPVNYAPNPDKHWIMRYTQPMKPIHMEFTNMLQRLQTLMSVDDSMETIYNMLVET 304
DB 305 H--GVYGD-----GQGLVK--GKSMPEYBPDIRVPFVVRGPNV-E 351
QY 305 GELDNTIYVVTADHGVHI-----GQGLVK--GKSMPEYBPDIRVPFVVRGPNV-E 351
DB 330 RLANDLTFTSDQGAHVEVSKGEIHGGSGNGIYVGRKANWEGGIRVGLLRWPRVIQ 389
QY 352 AGCLNPHVILNDLAPTLIDIAGLDIPAD--MDGKSLKLLD--TERPVN----- 397
DB 390 AGQKIDEPSTNMDIFPTVAKLAGAPLPEBRIIDGRDLMLLEGKQSRSDHEFLPHYCNAY 449
QY 398 ----RFLKKKQWRVDRSFLVER-----GKLLHKD-----NDKVDA 430
DB 450 LNAVRWHPQNSTSIWKAFFFTFNPNVGSNGCFATHVCFGSYVTHDPPLLEFDISKDP 509
QY 431 QEENFL-----PKYQKVLDLCORAE-----YOTACEQLGQK 461
DB 510 RERNPLTPASEPRPYEILKVMQEAARDHTQTLPPEVDQFSMNFLWKPMQLQCCPSTGLS 569
QY 462 WQCVEDATGK 471
DB 570 CQCDREKQDK 579

RESULT 9
ARS_STRPU STANDARD; PRT; 567 AA.
AC P50473;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Arylsulfatase precursor (EC 3.1.6.1) (Aryl-sulfate sulphohydrolase)
DE (ARS).
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC Strongylocentrotus.
OX NCBI_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89357267; PubMed=2767335;
RA Yang Q., Angerer L.M., Angerer R.C.;
RT "Structure and tissue-specific developmental expression of a sea
RT urchin arylsulfatase gene.";
RL Dev. Biol. 135:53-65(1989).
CC -!- CATALYTIC ACTIVITY: A phenol sulfate + H(2)O = a phenol + sulfate.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: EXPRESSION IS CONFINED TO ABORAL ECTODERM
CC CELLS AND THEIR PRECURSORS.
CC -!- DEVELOPMENTAL STAGE: LOW LEVELS ARE FOUND AT MESENCHYME BLASTULA
CC STAGE (24 HR), LEVELS INCREASE BY LATE GASTRULA STAGE AND ARE
CC MAINTAINED AT PLUTEUS STAGE.
CC -!- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M28404; AAA30036.1; .
CC PIR; A37362; A37362.
CC HSSP; P15289; IAUK.
CC InterPro; IPR000917; Sulfatase.
CC Pfam; PF00884; Sulfatase; 1.

```

```

DR PROSITE; PS00523; SULFATASE 1; 1.
DR PROSITE; PS00149; SULFATASE 2; 1.
KW Hydrolase; Signal; Glycoprotein.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 567 ARYL SULFATASE.
FT MOD_RES 115 115 2-AMINO-3-OXOPROPIONIC ACID (BY
FT SIMILARITY).
FT ACT_SITE 173 173 POTENTIAL.
FT CARBOHYD 179 179 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 228 228 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 542 542 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 567 AA; 62477 MW; D5B627993A4C4D6 CRC64;

Query Match 5.0%; Score 238; DB 1; Length 567;
Best Local Similarity 23.4%; Pred. No. 1.4e-09;
Matches 103; Conservative 58; Mismatches 167; Indels 112; Gaps 18;

QY 2 GPSSLVLCILSATVFLSGSSAFSLSHRLKGRFORDRENIRPNILVLTDDQDVGLGSM 61
DB 27 GPDABSLASLDRTATRRYGDGEDLL--HLGGTQGHRTAMTKPNVILLADDMGV--GDL 82
QY 62 QVANKTRR-----IMEQGTHTFINAFVTTMCCPSRSSILTKYVHNHNTYTNN----- 111
DB 83 SVYGHPTQBPFGFDQMANQGLRFTQGYSGDSVCTPSRSAIVTGRQPIRTGVYGEERIFLP 142
QY 112 --NCSSPSWQAQHSRTFAVYLNSTGYRTAPFGK--LNE---YNGSYVP----- 154
DB 143 WTTTGLPLYEV-----TIAEAMKAGAYTTGMVGMKHLGINENSSDGAHLPANRGDFVFG 197
QY 155 ---PGWKBN---VGLLKN-----SRFYNTLCRNGVKKEKHSYDSDYLTDLITNDS 200
DB 198 HNLPGNSWRCDTGLHQPDPDTNACFLYNSVTAQPFQHKG-----LTQLLRDDT 249
QY 201 VSFRTSKMYPHRPVLVMSHAAPHGPDSPAPQYSLRFPNASQHIPTSYNAPNPKHW 260
DB 250 VGFIEDN---VNKPFMYVVSFAHME-----TSLFSSD----- 278
QY 261 IMEYTGPMKPIHMEFTNMLQRL--QTLMSVDDSMETIYNMLVETGEIDNTYIVVTAHGG 319
DB 279 -----DFCTSRGRYGDNLREMQALISQIVTLVDNDIDNTVFTISHG 325
QY 320 YH---IGQFLVK---GKSMPEYBPDIRVPFVVRGPNVEAGCLNPHVILNDLAPTLIDI 372
DB 326 PHREYCGEGGDANVFRGKGQSWEGGHRIPYIVYVPGTISPGVSHEIVTSMIDIATAVNL 385
QY 373 AGLDIPAD--MDGKSLKLL 390
DB 386 GGSQLPTRIDYDGKCLKSVL 405

RESULT 10
IDS_HUMAN STANDARD; PRT; 550 AA.
ID -IDS_HUMAN
AC P22304; Q14604;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Iduronate 2-sulfatase precursor (EC 3.1.6.13).
GN IDS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 34-58 AND 456-473.
RC TISSUE=Endothelial cells;
RX MEDLINE=91046030; PubMed=2122463;
RA Wilson P.J., Morris C.P., Anson D.S., Occhiodoro T., Bielicki J.,
RA Clements P.R., Hopwood J.J.;
RT "Hunter syndrome: isolation of an iduronate-2-sulfatase cDNA clone
RT and analysis of patient DNA.";
RN Proc. Natl. Acad. Sci. U.S.A. 87:8531-8535(1990).
RN [2]

```



RP SEQUENCE FROM N.A.  
RX MEDLINE=94063929; PubMed=8244397;  
RA Wilson P.J., Meaney C.A., Hopwood J.J., Morris C.P.;  
RT "Sequence of the human iduronate-2-sulfatase (IDS) gene.";  
RL Genomics 17:773-775(1993).  
[3]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=96352905; PubMed=8717057;  
RX Timms K.M., Lu F., Shen Y., Pierson C.A., Muzny D.M., Gu Y.,  
RA Nelson D.L., Gibbs R.A.;  
RT "130 kb of DNA sequence reveals two new genes and a regional  
RT duplication distal to the human iduronate-2-sulfate sulfatase  
RT locus.";  
RL Genome Res. 5:71-78(1995).  
[4]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=96079126; PubMed=8530090;  
RX Malmgren H., Carlberg B.M., Pettersson U., Bondeson M.L.;  
RA "Identification of an alternative transcript from the human iduronate-  
RT 2-sulfatase (IDS) gene.";  
RL Genomics 29:291-293(1995).  
[6]  
RN SEQUENCE OF 1-398 FROM N.A.  
RP MEDLINE=93258349; PubMed=8490623;  
RX Floren R.H., Green E.P., Green P.M., Bentley D.R., Giannelli F.;  
RA "Determination of the organisation of coding sequences within the  
RT iduronate sulphate sulphatase (IDS) gene.";  
RL Hum. Mol. Genet. 2:5-10(1993).  
[7]  
RN REVIEW ON MPS-II VARIANTS.  
RP MEDLINE=94154729; PubMed=8111411;  
RX Hopwood J.J., Bunge S., Morris C.P., Wilson P.J., Steglich C.,  
RA Beck M., Schwinger E., Gal A.;  
RT "Molecular basis of mucopolysaccharidosis type II: mutations in the  
RT iduronate-2-sulphatase gene.";  
RL Hum. Mutat. 2:435-442(1993).  
[8]  
RN VARIANTS MPS-II ARG-135 AND GLY-422.  
RP MEDLINE=93265059; PubMed=1303211;  
RX Bunge S., Steglich C., Beck M., Rosenkranz W., Schwinger E.,  
RA Hopwood J.J., Gal A.;  
RT "Mutation analysis of the iduronate-2-sulfatase gene in patients with  
RT mucopolysaccharidosis type II (Hunter syndrome).";  
RL Hum. Mol. Genet. 1:335-339(1992).  
[9]  
RN VARIANTS MPS-II TRP-468.  
RX MEDLINE=93258320; PubMed=1284597;  
RA Crotti P.L., Bunge S., Anderson R.A., Whitley C.B.;  
RT "Mutation R468W of the iduronate-2-sulfatase gene in mild Hunter  
RT syndrome (mucopolysaccharidosis type II) confirmed by in vitro  
RT mutagenesis and expression.";  
RL Hum. Mol. Genet. 1:755-757(1992).  
[10]  
RN VARIANTS MPS-II ARG-86; ASP-94; ARG-120; PRO-221 AND GLY-422.  
RX MEDLINE=94108441; PubMed=8281149;  
RA Bunge S., Steglich C., Zuther C., Beck M., Morris C.P., Schwinger E.,  
RA Schinzel A., Hopwood J.J., Gal A.;  
RT "Iduronate-2-sulfatase gene mutations in 16 patients with  
RT mucopolysaccharidosis type II (Hunter syndrome).";  
RL Hum. Mol. Genet. 2:1871-1875(1993).  
[11]  
RN VARIANTS MPS-II GLU-68; HIS-293; GLY-478 AND ARG-485.  
RX MEDLINE=95072615; PubMed=7981716;  
RA Schroeder W., Wulff K., Wennert M., Seidlitz G., Herrmann F.H.;  
RT "Mutations of the iduronate-2-sulfatase (IDS) gene in patients with  
RT Hunter syndrome (mucopolysaccharidosis II).";  
RL Hum. Mutat. 4:128-131(1994).  
[12]  
RN VARIANTS MPS-II TRP-132; TYR-229; ARG-358; HIS-469 AND CYS-523.  
RX MEDLINE=95193786; PubMed=7887413;  
RA Jonsson J.J., Aronovich E.L., Braun S.E., Whitley C.B.;  
RT "Molecular diagnosis of mucopolysaccharidosis type II (Hunter  
RT syndrome) by automated sequencing and computer-assisted  
RT interpretation: toward mutation mapping of the iduronate-2-sulfatase  
RT gene.";  
RL Am. J. Hum. Genet. 56:597-607(1995).  
[13]  
RN VARIANTS MPS-II LEU-86; ASN-87; PRO-92; CYS-345 AND TRP-468.  
RX MEDLINE=95245347; PubMed=7728156;  
RA Popowska E., Rathmann M., Tytki-Szymanska A., Bunge S., Steglich C.,  
RA Schwinger E., Gal A.;  
RT "Mutations of the iduronate-2-sulfatase gene in 12 Polish patients  
RT with mucopolysaccharidosis type II (Hunter syndrome).";  
RL Hum. Mutat. 5:97-100(1995).  
[14]  
RN VARIANT MPS-II VAL-346.  
RX MEDLINE=95322987; PubMed=7599640;  
RA Li P., Huffman P., Thompson J.N.;  
RT "Mutations of the iduronate-2-sulfatase gene on a T146T background in  
RT three patients with Hunter syndrome.";  
RL Hum. Mutat. 5:272-274(1995).  
[15]  
RN VARIANTS MPS-II.  
RX MEDLINE=97094177; PubMed=8940265;  
RA Rathmann M., Bunge S., Beck M., Kresse H., Tytki-Szymanska A., Gal A.;  
RT "Mucopolysaccharidosis type II (Hunter syndrome): mutation 'hot spots'  
RT in the iduronate-2-sulfatase gene.";  
RL Am. J. Hum. Genet. 59:1202-1209(1996).  
[16]  
RN VARIANTS MPS-II LEU-333 AND ASP-346.  
RX MEDLINE=96163494; PubMed=8566953;  
RA Olsen T.C., Eiken H.G., Knappekog P.M., Kase B.F., Mansson J.-E.,  
RA Boman H., Apold J.;  
RT "Mutations in the iduronate-2-sulfatase gene in five Norwegians with  
RT Hunter syndrome.";  
RL Hum. Genet. 97:198-203(1996).  
[17]  
RN VARIANTS MPS-II LEU-333 AND GLY-334.  
RX MEDLINE=96255081; PubMed=8830188;  
RA Li P., Thompson J.N.;  
RT "Detection of four novel mutations in the iduronate-2-sulphatase gene  
RT by single-strand conformation polymorphism analysis of genomic  
RT amplicons.";  
RL J. Inher. Metab. Dis. 19:93-94(1996).  
[18]  
RN VARIANTS MPS-II ASP-63; THR-347; GLN-468 AND LEU-468.  
RX MEDLINE=97365936; PubMed=9222763;  
RA Villani G.R.D., Balzano N., Grosso M., Salvatore F., Izzo P.,  
RA di Natale P.;  
RT "Mucopolysaccharidosis type II: identification of six novel mutations  
RT in Italian patients.";  
RL Hum. Mutat. 10:71-75(1997).  
[19]  
RN VARIANT MPS-II GLN-468.  
RX MEDLINE=98041699; PubMed=9375851;  
RA Sukegawa K., Song X.-Q., Masuno M., Fukao T., Shimozaawa N., Fukuda S.,  
RA Isozaki K., Nishio H., Matsuo M., Tomatsu S., Kondo N., Orii T.;  
RT "Hunter disease in a girl caused by R468Q mutation in the  
RT iduronate-2-sulfatase gene and skewed inactivation of the X  
RT chromosome carrying the normal allele.";  
RL Hum. Mutat. 10:361-367(1997).  
[20]  
RN VARIANTS MPS-II ASN-45; TYR-115; LEU-228; ARG-266; LYS-434; LYS-485  
RP AND CYS-502.  
RX MEDLINE=99092178; PubMed=9875019;  
RA Vafiadaki E., Cooper A., Heptinstall L.E., Hatton C.E., Thornley M.,  
RA Wraith J.E.;

"Mutation analysis in 57 unrelated patients with MPS II.";  
 Arch. Dis. Child. 79:237-241(1998).  
 [21]  
 RT VARIANTS MPS-II.  
 RA MEDLINE=9911886; PubMed=9921913;  
 RA Karsten S., Voskobojeva E., Tishkanina S., Pettersson U.,  
 RA Krasnopolskaja X., Bondeson M.-L.;  
 RT "Mutational spectrum of the iduronate-2-sulfatase (IDS) gene in 36  
 RT unrelated Russian MPS II patients.";  
 RL Hum. Genet. 103:732-735(1998).  
 [22]  
 RP VARIANTS MPS-II LEU-86; HIS-88; PRO-88; ILE-118 AND HIS-266.  
 RA Balzano N., Villani G.R.D., Grosso M., Izzo P., di Natale P.;  
 RT "Detection of four novel mutations in the iduronate-2-sulfatase  
 RT gene.";  
 RL Hum. Mutat. 11:333-333(1998).  
 [23]  
 RP VARIANTS MPS-II THR-85; HIS-88; ILE-349 AND VAL-521.  
 RA MEDLINE=98112423; PubMed=9452044;  
 RA Gort L., Coll M.J., Chabas A.;  
 RT "Mutations in the iduronate-2-sulfatase gene in 12 Spanish patients  
 RT with Hunter disease.";  
 RL Hum. Mutat. Suppl. 1:S66-S68(1998).  
 [24]  
 RP VARIANTS MPS-II PHE-143; TRP-184; VAL-269 AND HIS-348.  
 RA Karsten S.L., Voskobojeva E., Carlberg B.-M., Kleijer W.J.,  
 RA Toennesen T., Pettersson U., Bondeson M.-L.;  
 RT "Identification of 9 novel gene mutations in 19 unrelated Hunter  
 RT syndrome (Mucopolysaccharidosis type II) patients.";  
 RL Hum. Mutat. 12:433-433(1998).  
 [25]  
 RP VARIANTS MPS-II PHE-143; LYS-341; TYR-342 AND PHE-491.  
 RA MEDLINE=99235558; PubMed=10220152;

Query Match 4.9%; Score 234.5; DB 1; Length 550;  
 Best Local Similarity 21.2%; Pred. No. 2.4e-09;  
 Matches 114; Conservative 80; Mismatches 189; Indels 155; Gaps 23;  
 QY 1 MCPPS-----LVCLLSATVFSLLGSSAFSLSHRLKGRFQDRNRNIRPNILVLTDDQD 55  
 DB 1 MPPPTGRGLLGLVLSVVCVALGSET-----QANSTTDALNVLIIIVDDL 48  
 QY 56 VELGSMQVNNKTRIMEQGTTH---FINAVTTPMCCPSRSILTGK-----YVHNHTY 107  
 DB 49 PSLGCVGDKLVSPNIDQLASHSLFLQNAFAQAVCAPSVSFLTRPDTIRLYDNFSY 108  
 QY 108 TNENCSPPSQAQHS-RTFAVYLNSTGYRTAFPKYL-----NBYNGSVYP- 154  
 DB 109 -----NRVHAGNFSTIPQYFKNGYVTVMSYGVKFHEGIGSSNHTDDSPYSNPF 158  
 QY 155 -PGWKEWGLLNKSNRYNYTLCRNGYKKGHS-----DYSKDYLD-LITNDSVSFF 204  
 DB 159 HP-----SSEKYENTKTCRGPDELHANLLCPVDVLPDPEGTLPDKOSTEQAIQL 209  
 QY 205 RTSKMYPHRPVLMVISHAHPGSDPAQYSRLPFNASQHTIPSYNAPND----- 257  
 DB 210 ERKMT--SASPFFLAVGVHKPHIPRYPKFKQLYP--LENIT---LAPDPEVPDGLPP 261  
 QY 258 ---KHWI-----MRYTGPMKPIHEFTNMLQRKELQTLMSVDDSMETIYNMLVET 304  
 DB 262 VAYNPMMDIQRQEDVOALNINIPYGPDPVDFORKIRQSVFASVLDTVQVRLLSALDDL 321  
 QY 305 GELDNTIYVYADGHYGHIGFGLVNGKSPYBFDIRVP--FYVRGN----- 349  
 DB 322 QLANSTIIAFTSDHGWALGEHG-ERAKTSNFDVATHVPLIFVFGRTASLPEAGEKLFPY 380  
 QY 350 -----VRAGCLNPHIVLNIDLAFTILDIDPADMDGKSIKLKLDTERPVNRF 399  
 DB 381 LDPFDSASQLMEFGQSMDELVSFLPTLAGLAGLQVP-----PRCPVPSF 427  
 QY 400 HLKCKKRVNRDVSFLVERGKLLKHKRNDKVDAAEENFLP-----KYGRVKDLQ 447  
 DB 428 HVE-----LCREGKLLKHFRFR-DLBEDPYLPGNPRELIAYSOYRPSDIPQ 474

RESULT 11  
 ARS HEMPU STANDARD; PRT; 551 AA.  
 AC P14000; 1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Arylsulfatase precursor (BC 3.1.6.1) (Aryl-sulfate sulphohydrolase)  
 DE (ARS).  
 OS Hemocentrotus pulcherrimus (Sea urchin).  
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
 OC Echinoidea; Euechinoidea; Echinacea; Echinacea; Strongylocentrotidae;  
 OC Hemocentrotus.  
 OX NCBI\_TaxID=7650;  
 [1]  
 RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Pluteus.  
 RX MEDLINE=89030659; PubMed=3181160;  
 RA Sasaki H., Yamada K., Akasaka H., Suzuki K., Saito A., Sato M.,  
 RA Shimada H.;  
 RT "cDNA cloning, nucleotide sequence and expression of the gene for  
 RT arylsulfatase in the sea urchin (Hemocentrotus pulcherrimus)  
 RT embryo.";  
 RL Eur. J. Biochem. 177:9-13(1988).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=90092130; PubMed=2598936;  
 RA Yamada K., Akasaka K., Shimada H.;  
 RT "Structure of sea-urchin arylsulfatase gene.";  
 RL Eur. J. Biochem. 186:405-410(1989).  
 CC -!- FUNCTION: MAY BE A STRUCTURAL COMPONENT OF THE EXTRACELLULAR  
 CC MATRICES INVOLVED IN CELL MOVEMENT DURING MORPHOGENESIS.  
 CC -!- CATALYTIC ACTIVITY: A phenol sulfate + H(2)O = a phenol + sulfate.  
 CC -!- SUBCELLULAR LOCATION: IN BOTH THE CYTOPLASM AND THE EXTRACELLULAR  
 CC MATRICES.  
 CC -!- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; X17015; CAA34881.1; -.  
 DR PIR; S01793; S01793.  
 DR HSPP; P15289; LAUK.  
 DR InterPro; IPR000917; Sulfatase.  
 DR Pfam; PF00884; Sulfatase; 1.  
 DR PROSITE; PS00523; SULFATASE\_1; 1.  
 DR PROSITE; PS00149; SULFATASE\_2; 1.  
 KW Hydrolyase; Signal; Glycoprotein; Extracellular matrix.  
 FT SIGNAL 1 20  
 FT CHAIN 21 551 ARYLSULFATASE.  
 FT MOD\_RES 21 21 BLOCKED.  
 FT MOD\_RES 100 100 2-AMINO-3-OXOPROPIONIC ACID (BY  
 FT SIMILARITY).  
 FT ACT\_SITE 158 158 POTENTIAL.  
 FT CARBOHYD 164 164 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 213 213 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 296 296 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 551 AA; 60952 MW; 54C1AAC14D6710C9 CRC64;  
 Query Match 4.9%; Score 232; DB 1; Length 551;  
 Best Local Similarity 23.6%; Pred. No. 3.7e-09;  
 Matches 105; Conservative 63; Mismatches 161; Indels 116; Gaps 21;  
 QY 42 IRPNILVLTDDQDVELGSMQVNNKTRIMEQ-----GTHFINAVTTPMCCPSRS 93  
 DB 50 VKPNVLLVAD-----HMGSGDLTSYGHGTQEGAFIDKMAEGLRFTNGVGVDAVCTPSRS 105

QY 94 SILTKYVHNNTYNNENCSPSHQ---A0HESRTFAVYLNSTGYRTAFGKY---LNE 147  
 DB 106 AIMGRLPVRIGTF--GETRVFLPMTKGLPKSELTTAEAMKEAGYATGMVCKWHLGNE 163  
 QY 148 ---YNGSVYP-----PGKWE-----VGLLK-----NSRFY---NYTLCRNGVKE 181  
 DB 164 NSSTDGALPHNPGDFVGHNLPTFNSWCDTGLHKDFPDSPQCYLYVNAVTLV----- 217  
 QY 182 KHGSDYSKDYLTDLITNDSVFFRTSKQYRHPVLMVISHAAPHGPDSPAPOYSRLFPN 241  
 DB 218 --SOPYQHKGLTQLFTDDALGFI-----EDN----- 241  
 QY 242 ASOHTPSVNYAPNDKWIWRYGPMKPIHMEFTNMLQKRL-QLTMSVDDSMETIYNM 300  
 DB 242 ---HADPFYVAFAMHTSLFSSD-----DFCSSTRGKYGDNLLEHMDAVKQIVDK 291  
 QY 301 LVETGELONTYIVYADHGYHI-----GQGLVK-GKSMPEYFDIRVPYVRGPNVEAG 353  
 DB 292 LEENNISENTIIFISDHGPHREYCEEGDASIFRGKSHSWEGHRIPIYIVYWGCTISP 351  
 QY 354 CUNPHVINIDILPILDIAGLIDIPAD--MDGKSILKLL--DTERPVNRFLKXKQWVR 409  
 DB 352 GISNEIVTSMIDIAADLGGTTLPTDRIYDGKSIKDVLEGASPHSSFFYCK----- 406  
 QY 410 DSFLVER-GKLLHKEDNDKVDQBE 433  
 DB 407 UNLMAVRVGKYKAHFTQTVRSQDE 431

RESULT 12  
 IDS\_MOUSE  
 ID\_IDS\_MOUSE STANDARD; PRT; 563 AA.  
 AC Q08890;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Iduronate 2-sulfatase precursor (EC 3.1.6.13).  
 GN IDS.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Thymus;  
 RX MEDLINE=93315172; PubMed=8325651;  
 RA Daniele A., Faust C.J., Herman G.E., di Natale P., Ballabio A.;  
 RT "Cloning and characterization of the cDNA for the murine iduronate  
 sulfatase gene.";  
 RL Genomics 16:755-757(1993).  
 CC -!- FUNCTION: REQUIRED FOR THE LYSOSOMAL DEGRADATION OF HEPARAN  
 CC SULFATE AND DERMATAN SULFATE.  
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 2-sulfate groups of the L-  
 CC iduronate 2-sulfate units of dermatan sulfate, heparan sulfate and  
 CC heparin.  
 CC -!- SUBCELLULAR LOCATION: Lysosomal.  
 CC -!- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; L07921; AAA37880.1; -  
 DR PIR; A47153; A47153.  
 DR MGD; MGI:96417; Ids.  
 DR InterPro; IPR000917; Sulfatase.  
 DR Pfam; PF00884; Sulfatase; 1.  
 DR PROSITE; PS00523; SULFATASE\_1; 1.

DR PROSITE; PS00149; SULFATASE\_2; 1.  
 KW Hydrolase; Glycoprotein; Lysosome; Zymogen; Signal.  
 FT SIGNAL 1 ? POTENTIAL.  
 FT PROPEP 7 59 BY SIMILARITY.  
 FT CHAIN 60 563 IDURONATE 2-SULFATASE.  
 FT MOD\_RES 99 99 2-AMINO-3-OXOPROPIONIC ACID (BY  
 FT SIMILARITY).  
 FT CARBOHYD 130 130 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 159 159 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 261 261 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 526 526 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 550 550 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 563 AA; 63437 MW; 52D008F91BA88FAC CRC64;  
 Query Match 4.8%; Score 226; DB 1; Length 563;  
 Best Local Similarity 21.7%; Pred. No. 1e-08;  
 Matches 101; Conservative 69; Mismatches 199; Indels 96; Gaps 17;  
 QY 16 FSLGSSNAPLSHRLKGRFQDRNIRPNILVLTDDQVELGSMQVMKTRIMEQGG 75  
 DB 28 FSLGSSFCIALESAAQGSATDAL---NILLIIVDDLRESLGCYGLKVRSPNIDOLA 83  
 QY 76 TH---FINAVFTTPMCCPSRSILTKG---YVNHNTYNNENCSPSHQAQHS-RT 126  
 DB 84 SHSVLFQNAFAQAVCAPSRVSFLTGRPDITRLYDENS- ---WRVHSGNFST 133  
 QY 127 PAVYLNSTGYRTAFGKYL-----NEYNGSVYPPGKWEVGLLKNRFPNYTLCRNGVKE 181  
 DB 134 IPOYFKNGYVTMSVGKVFHFGIISNHSDDY-PYSWSFPYPHPSSEKYENTKCKGQDGK 192  
 QY 182 KHG-----SDYSKDYLTDLITNDSVSFFRTSKQYRHPVLMVISHAAPHGPDSPA 232  
 DB 193 LHANLLCPVDVADVPEGLTPDKQSTEEA--IRLLEKMTSGSPFFLAVGYHKKPHIPRYP 250  
 QY 233 PQYSRLFPNASCHITPSYNYAPNPKH-----WI-----MRYTGMPKP 270  
 DB 251 KEFQKLYP--LENITLA-----PDHPVDSLPVYVNPMDIREREDVQALNISVPYGP 302  
 QY 271 IHMEFTNMLQKRLQTLMSVDDSMETIYNMLVETGELONTYIVYADHGYHIGQGP----- 326  
 DB 303 IPEDFQRIQSYFASVSYLDTQVGHVLSALDDRLAINTIAFTSDHGWALGHEGWAK 362  
 QY 327 -----LVKGSMPYE-----FDIRVPYVRGPNVEAGCLNPHIVLINIDLA 366  
 DB 363 YSNFEDVATVPLMLYVPGRTAPLPAAGOKLFPYRDPDPASDMDAGRHTEDLVLSLF 422  
 QY 367 PTLIDLAGLIDPADMDGKSILKLLDTERPVNRFLKXKQWVRDS 411  
 DB 423 PTLAGLAGL--EVLGAPSLPLMLSLFAEKARIFRSICSSMTWKR 465

RESULT 13  
 GACS\_HUMAN  
 ID\_GACS\_HUMAN STANDARD; PRT; 522 AA.  
 AC P34059;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE N-acetylgalactosamine-6-sulfatase precursor (EC 3.1.6.4) (N-  
 DE acetylgalactosamine-6-sulfate sulfatase) (Galactose-6-sulfate  
 DE sulfatase) (GALNAC6S sulfatase) (Chondroitinsulfatase)  
 DE (Chondroitinase).  
 GN GALNS.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Placenta;  
 RX MEDLINE=92095973; PubMed=1755850;  
 RX Tomatsu S., Fukuda S., Masue M., Sukegawa K., Fukao T., Yamagishi A.,



RESULT 14  
ARSA\_MOUS  
ID ARSA  
AC P504

QY 44 PNILLVLTDQDV-ELGSMQVNVKTRIMEQ---CGTHFINAFVTPMCCSRSSILTKG 99  
 DB 20 PNILLFADDLGGLGYGHPSTTPNLDQALAEGLRFTDFYVPSLCTPSRAALLTGR 79  
 QY 100 YVHNHTYTNNECSSPSWQA--QHESTFAVYLNSTGRTAFKYLNEY--NGSVVPP 155  
 DB 80 LPVRSAMY---PGVLGSPSQGLPLEELTLAEVLAARGYLTMAGKWHLGVPGEAFLPP 136  
 QY 156 --GWKEWVGLLNKSNRYNLTGRN-----GVKERHGSYDXYLTDITNDVSFFRTS 207  
 DB 137 HQGFHRLFGIPYS---HDQGPCNLTCFPPDIPCKGGCDQGLVPI--PULAN-----LTV 186  
 QY 208 KQMYPHRPVLMVISHAAPHGPDSPQVSR-LFPNASQHIPTPSYNYAPNPKHIMRYTG 266  
 DB 187 EAQPPWLP-----GLEARYVSFRDLMAAQGRGPFYLYASHHTY--POFSG 234  
 QY 267 PMKPIMHFTNNLQRL--QTLSMVDSDNETIYNMLVETGELDNITYIVYTADHGVI--- 322  
 DB 235 -----QSFTKESGRGPGCDLSMELDGAVALMTTVDGLGLEETLVIFTADNGFELMRM 288  
 QY 323 ---GQFGLVK--GKSNPYEFDIRVPYVGRGNVACGLNPHVNLIDAPTILDIAGLDIP 378  
 DB 289 SNGGCSGLLRCKGKTTFEGGVRPALVMPGHITPGVTHELASSLDLPTLAALTGAPLP 348  
 QY 379 -ADMCKSILKLL 390  
 DB 349 NVTLGQVLSPLL 361

## RESULT 15

ARSB HUMAN  
 ID ARSB HUMAN STANDARD; PRT; 533 AA.  
 AC P15848;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DE Arylsulfatase B precursor (EC 3.1.6.12) (ASB) (N-acetylgalactosamine-4-sulfatase) (G4S).  
 GN ARSB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI TaxID:9606;  
 [1] SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE=90153594; PubMed=2303452;  
 RA Peters C., Schmidt B., Rommerskirch W., Rupp K., Zuehlendorf M., Vingron M., Meyer H.E., Pohlmann R., von Figura K.;  
 RT "Phylogenetic conservation of arylsulfatases. cDNA cloning and expression of human arylsulfatase B.";  
 RL J. Biol. Chem. 265:3374-3381(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90152677; PubMed=1968043;  
 RA Schuchman E.H., Jackson C.E., Desnick R.J.;  
 RT "Human arylsulfatase B: MOPAC cloning, nucleotide sequence of a full-length cDNA, and regions of amino acid identity with arylsulfatases A and C.";  
 RL Genomics 6:149-158(1990).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93332648; PubMed=7687847;  
 RA Modaresi S., Rupp K., von Figura K., Peters C.;  
 RT "Structure of the human arylsulfatase B gene.";  
 RL Biol. Chem. Hoppe-Seyler 374:327-335(1993).  
 RN [4]  
 RP SEQUENCE OF 1-104 FROM N.A.  
 RX MEDLINE=92028992; PubMed=1930244;  
 RA Litjens T., Morris C.P., Gibson G.J., Beckmann K.R., Hopwood J.J.;  
 RT "Human N-acetylgalactosamine-4-sulphatase: protein maturation and isolation of genomic clones.";  
 RL Biochem. Int. 24:209-215(1991).

RN [5]  
 RP 2-AMINO-3-OKOPROPIONIC ACID MODIFICATION SITE.  
 RX MEDLINE=95354208; PubMed=7628016;  
 RA Schmidt B., Selmer T., Ingendoh A., von Figura K.;  
 RT "A novel amino acid modification in sulfatases that is defective in multiple sulfatase deficiency.";  
 RL Cell 82:271-278(1995).  
 RN [6]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).  
 RX MEDLINE=971184692; PubMed=9032078;  
 RA Bond C.S., Clements P.R., Ashby S.J., Collyer C.A., Harrop S.J., Hopwood J.J., Guss J.M.;  
 RT "Structure of a human lysosomal sulfatase.";  
 RL Structure 5:277-289(1997).  
 RN [7]  
 RP VARIANTS MPS-VI ARG-117; PRO-236 AND TYR-405.  
 RX MEDLINE=92197625; PubMed=1550123;  
 RA Jin W.-D., Jackson C.E., Desnick R.J., Schuchman E.H.;  
 RT "Mucopolysaccharidosis type VI: identification of three mutations in the arylsulfatase B gene of patients with the severe and mild phenotypes provides molecular evidence for genetic heterogeneity.";  
 RL Am. J. Hum. Genet. 50:795-800(1992).  
 RN [8]  
 RP VARIANT MPS-VI VAL-137, AND VARIANT MET-376.  
 RX MEDLINE=92042029; PubMed=1718978;  
 RA Wicker G., Prill V., Brooks D., Gibson G., Hopwood J., von Figura K., Peters C.;  
 RT "Mucopolysaccharidosis VI (Maroteaux-Lamy syndrome). An intermediate clinical phenotype caused by substitution of valine for glycine at position 137 of arylsulfatase B.";  
 RL J. Biol. Chem. 266:21386-21391(1991).  
 RN [9]  
 RP VARIANTS MPS-VI MET-92; GLN-95; CYS-210; PRO-393 AND PRO-498.  
 RX MEDLINE=96213747; PubMed=8651289;  
 RA Litjens T., Brooks D.A., Peters C., Gibson G.J., Hopwood J.J.;  
 RT "Identification, expression, and biochemical characterization of N-acetylgalactosamine-4-sulfatase mutations and relationship with clinical phenotype in MPS-VI patients.";  
 RL Am. J. Hum. Genet. 58:1127-1134(1996).  
 RN [10]  
 RP VARIANTS MPS-VI TRP-152 AND GLN-160.  
 RX MEDLINE=94171224; PubMed=8125475;  
 RA Voskoboeva E., Isbrandt D., von Figura K., Krasnopol'skaya X., Peters C.;  
 RT "Four novel mutant alleles of the arylsulfatase B gene in two patients with intermediate form of mucopolysaccharidosis VI (Maroteaux-Lamy syndrome).";  
 RL Hum. Genet. 93:259-264(1994).  
 RN [11]  
 RP VARIANT MPS-IV ARG-302.  
 RA Villani G.R.D., Balzano N., di Natale P.;  
 RT "Two novel mutations of the arylsulfatase B gene in two Italian patients with severe form of mucopolysaccharidosis.";  
 RL Hum. Mutat. 11:410-410(1998).  
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of the 4-sulfate groups of the N-acetyl-D-galactosamine 4-sulfate units of chondroitin sulfate and dermatan sulfate.  
 CC -1- SUBUNIT: Monomer.  
 CC -1- SUBCELLULAR LOCATION: Lysosomal.  
 CC -1- DISEASE: DEFECTS IN ARSB ARE THE CAUSE OF MAROTEAUX-LAMY SYNDROME; ALSO KNOWN AS MUCOPOLYSACCHARIDOSIS TYPE VI (MPS-VI). THIS DISEASE IS CHARACTERIZED BY THE ACCUMULATION OF DERMATAN SULFATE IN LYSOSOMES. CLINICAL FEATURES CAN INCLUDE ABNORMAL GROWTH, SHORT STATURE, STIFF JOINTS, SKELETAL MALFORMATIONS, CORNEAL CLOUDING, HEPATOSPLENOMEGALY, AND CARDIAC ABNORMALITIES. A WIDE VARIATION IN CLINICAL SEVERITY IS OBSERVED.  
 CC -1- DISEASE: MULTIPLE SULFATASE DEFICIENCY (MSD) IS A DISORDER THAT COMBINES FEATURES OF METACHROMATIC LEUKODYSTROPHY AND OF MUCOPOLYSACCHARIDOSIS; IT IS CHARACTERIZED BY A DECREASED ACTIVITY OF ALL KNOWN SULFATASES. IT SEEMS TO BE CAUSED FROM THE LACK OF POST-TRANSLATIONAL MODIFICATION OF A CYSTEINE INTO 2-AMINO-3-OKOPROPIONIC ACID.  
 CC -1- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2003, 12:56:09 ; Search time 50 Seconds

(without alignments)  
2761.842 Million cell updates/sec

Title: US-09-970-287-2

Perfect score: 4751

Sequence: 1 MGPPSLVCLLSATVFSILG.....PEMKRPSKSLGQLWEGWEG 870

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

```

1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4751	100.0	870	23	ABB80921
2	4751	100.0	885	23	ABB80922
3	4746	99.9	870	22	AAB85774
4	4746	99.9	870	22	AAU79215
5	4746	99.9	870	22	AAE01440
6	4746	99.9	870	23	ABG71836
7	4746	99.9	870	23	ABG63903
8	4742	99.8	870	22	AAE01471
9	4742	99.8	870	23	ABG63902
10	4751	100.0	870	23	ABB80921
11	4751	100.0	885	23	ABB80922
12	4746	99.9	870	22	AAB85774
13	4746	99.9	870	22	AAU79215
14	4746	99.9	870	22	AAE01440
15	4746	99.9	870	23	ABG71836
16	4746	99.9	870	23	ABG63903
17	4742	99.8	870	22	AAE01471
18	4742	99.8	870	23	ABG63902

10	4742	99.8	870	23	ABB90378	Human polypeptide
11	4719.5	99.3	867	21	AAU66648	Membrane-bound pro
12	4719.5	99.3	867	22	AAU12181	Human PRO1120 poly
13	4719.5	99.3	867	22	AAB65171	Human PRO1120 (UNQ
14	4719.5	99.3	867	23	ABB95459	Human angiogenesis
15	4719.5	99.3	867	23	ABB84853	Human PRO1120 prot
16	4719.5	99.3	867	24	ABU66579	Human PRO polypept
17	4719.5	99.3	867	24	ABU66855	Human secreted/tra
18	4719.5	99.3	867	24	ABU59660	Novel secreted and
19	4719.5	99.3	867	24	ABU59064	Novel human secret
20	4719.5	99.3	867	24	ABU59211	Human secreted/tra
21	4719.5	99.3	867	24	ABU59360	Novel human secret
22	4719.5	99.3	867	24	ABU60495	Human secreted/tra
23	4719.5	99.3	867	24	ABU57986	Human PRO polypept
24	4719.5	99.3	867	24	ABU58917	Human secreted/tr
25	4719.5	99.3	867	24	ABU13877	Human PRO1120 poly
26	4719.5	99.3	867	24	ABU10832	Human PRO polypept
27	4612	97.1	850	22	AAU79216	Human protein SEQ
28	4469.5	94.1	875	23	ABG71838	Mouse SULF2 protei
29	4343	91.4	818	21	ABA42268	Human ORFX ORF2032
30	3524	74.2	643	22	AAU80199	Human protein SEQ
31	3524	74.2	643	22	AAU80200	Breast cancer prot
32	3175.5	66.8	871	21	AAB00191	Human protein SEQ
33	3175.5	66.8	871	22	AAB85481	Human 23553 sulfat
34	3175.5	66.8	871	23	ABG71835	Human SULF1 protei
35	3175.5	66.8	871	23	ABG69640	Human secreted pro
36	3175.5	66.8	871	24	ABR47506	Breast cancer asso
37	3117.5	65.6	867	22	AAE00434	Quail sulfatase (Q
38	3059	64.4	818	23	AAU69417	Lung small cell ca
39	3059	64.4	1611	22	AAE00438	Human sulfatase (H
40	2973	62.6	800	22	AAU25714	Human protein sequ
41	2696	56.7	490	22	AAU93919	Human polypeptide,
42	2592	56.7	490	22	AAU93864	Human polypeptide,
43	2352	45.5	434	22	AAU01522	Human gene 5 encod
44	2093	44.1	455	22	AAU00437	Mouse sulfatase (M
45	1714.5	36.1	1114	22	ABB71505	Drosophila melanog

#### ALIGNMENTS

```

RESULT 1
ABB80921
ID ABB80921 standard; Protein; 870 AA.
XX
AC ABB80921;
XX
DT 08-OCT-2002 (first entry)
XX
DE Human sulphatase polypeptide, 22437.
XX
KW Human; sulphatase; 22437; cytosolic; vulnary; neuroprotective;
XX gene therapy; enzyme.
XX
OS Homo sapiens.
XX
PN WO200252019-A2.
XX
PD 04-JUL-2002.
XX
PF 03-OCT-2001; 2001WO-US30856.
XX
PR 21-DEC-2000; 2000US-257082P.
XX
(PMILL-) MILLENNIUM PHARM INC.
XX
Glucksmann MA, Rudolph-Owen LA;
XX
WPI; 2002-566677/60.
XX
N-PSDB; ABB86352, ABB86353.
XX
Identifying agents for modulating (e.g. inhibiting) e.g. tumor
PT establishment, growth or metastases, neuron growth, or wound healing by

```





Db 16 MGPPSLVCLLSATVFSLLGGSSAFSLSHRLKGRFQDRRNIRPNIIILVLTDDQDELGS 75  
 QY 61 MQVMNKTTRIMEQGGTHFINAFVTPMCCPSRSSILTGKYVHNHTYTNNECSPSQWA 120  
 Db 76 MQVMNKTTRIMEQGGTHFINAFVTPMCCPSRSSILTGKYVHNHTYTNNECSPSQWA 135  
 QY 121 QHESRTFVYLINSTGYRTAFPGKYLYNEVNGSVPPGKWEVGLLKNRFRNYTLCRNGVK 180  
 Db 136 QHESRTFVYLINSTGYRTAFPGKYLYNEVNGSVPPGKWEVGLLKNRFRNYTLCRNGVK 195  
 QY 181 EXHGSYDYSKDYLTDLITNDSVSFFRTSKMTYPHRPLVLMVISHAAPHGPDSPAPQYSLRFP 240  
 Db 196 EXHGSYDYSKDYLTDLITNDSVSFFRTSKMTYPHRPLVLMVISHAAPHGPDSPAPQYSLRFP 255  
 QY 241 NASOHITPSYAPNDKHMIRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300  
 Db 256 NASOHITPSYAPNDKHMIRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 315  
 QY 301 LVETGELDNITYIVYTADHGYHIGFGLVKGKSMPEYFDIRVPFVYRGPNVEAGCLNPHIV 360  
 Db 316 LVETGELDNITYIVYTADHGYHIGFGLVKGKSMPEYFDIRVPFVYRGPNVEAGCLNPHIV 375  
 QY 361 LNIDLAPTILDIAGLDIPADMDGKSILKLLDTERPVRNPHLKKQVRVWDSFLVERGKLL 420  
 Db 376 LNIDLAPTILDIAGLDIPADMDGKSILKLLDTERPVRNPHLKKQVRVWDSFLVERGKLL 435  
 QY 421 HKRDNKVDQAEENFLPKYQVRKDLQCBAEYOTACEQLGQKQWQVEDATGKLLHKCKGP 480  
 Db 436 HKRDNKVDQAEENFLPKYQVRKDLQCBAEYOTACEQLGQKQWQVEDATGKLLHKCKGP 495  
 QY 481 MRLGGSRALSNLVPKYVQGSSEACTCDSDYKLSLAGRRKKLPKKYKASYVRSRSIRSV 540  
 Db 496 MRLGGSRALSNLVPKYVQGSSEACTCDSDYKLSLAGRRKKLPKKYKASYVRSRSIRSV 555  
 QY 541 AIEVDGRVYHVLGDAAPRNLTKRHWFGAPEDQDDKDGDPGSGTGLPDYSAANPIKVT 600  
 Db 556 AIEVDGRVYHVLGDAAPRNLTKRHWFGAPEDQDDKDGDPGSGTGLPDYSAANPIKVT 615  
 QY 601 HRCYILENDTVQCDLDLYKSLQAKWKHLHDHEITLQNKIKNLRVGRHLKKRPEEC 660  
 Db 616 HRCYILENDTVQCDLDLYKSLQAKWKHLHDHEITLQNKIKNLRVGRHLKKRPEEC 675  
 QY 661 DCHKISYHTQHKRLKRGSSLHPFRKGLQEKDKVLLREOKRKKLKLKRLQNNDDTC 720  
 Db 676 DCHKISYHTQHKRLKRGSSLHPFRKGLQEKDKVLLREOKRKKLKLKRLQNNDDTC 735  
 QY 721 SMPGLTCFTHDNQHWQTAFFWTILGPFCACTSANNTYWCMTINETHNFLECFEATGPLE 780  
 Db 736 SMPGLTCFTHDNQHWQTAFFWTILGPFCACTSANNTYWCMTINETHNFLECFEATGPLE 795  
 QY 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCGYKQCNPRTRNMDLGLKDGGSYE 840  
 Db 796 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCGYKQCNPRTRNMDLGLKDGGSYE 855  
 QY 841 QYRQPRKRWPEMKRPSKSLGQWEGWEG 870  
 Db 856 QYRQPRKRWPEMKRPSKSLGQWEGWEG 885

RESULT 3

AAB85774

ID AAB85774 standard; Protein; 870 AA.

XX

AAB85774;

XX

29-OCT-2001 (first entry)

XX

Human drug metabolizing enzyme (ID No. 1558210CD1).

DE

KW Drug metabolizing enzyme; DME; immunosuppressive; cytostatic; ophthalmic;  
 KW hepatocellular; antiallergic; antiaesthetic; antibacterial; antiviral;  
 KW antisense therapy; gene therapy; human.

XX

OS Homo sapiens.  
 XX WO200159127-A2.  
 XX 16-AUG-2001.  
 XX 08-FEB-2001; 2001WO-US04423.  
 XX 11-FEB-2000; 2000US-0181856.  
 PR 17-FEB-2000; 2000US-0183684.  
 PR 25-FEB-2000; 2000US-0185141.  
 PR 03-MAR-2000; 2000US-0186818.  
 PR 09-MAR-2000; 2000US-0188345.  
 PR 17-MAR-2000; 2000US-0189997.  
 XX (INCY-) INCYTE GENOMICS INC.  
 PA Tang YT, Yue H, Baughn MR, Yao MG, Bandman O, Azimzai Y, Lal P;  
 PI Gandhi AR, Ring HZ, Shih LL, Yang J, Policky JB;  
 XX WPI; 2001-514673/56.  
 DR N-PSDB; AAH76201.  
 XX Isolated polypeptide encoding a drug metabolizing enzyme useful for the  
 PT diagnosis, treatment, and prevention of autoimmune/inflammatory, cell  
 PT proliferative, developmental and endocrine disorders -  
 XX Claim 1; Page 133-134; 150pp; English.  
 PS The invention provides human drug metabolizing enzymes (DME) and  
 CC polynucleotides encoding the DMEs. The DME can be expressed by standard  
 CC recombinant methodology. DMEs and their agonists and antagonists are  
 CC useful for the diagnosis, treatment, and prevention of autoimmune/  
 CC inflammatory, cell proliferative, developmental, endocrine such as  
 CC aneurysm, eye, metabolic, and gastrointestinal disorders, including liver  
 CC disorders and infection. The present sequence represents a human DME.  
 XX SQ Sequence 870 AA;

Query Match 99.9%; Score 4746; DB 22; Length 870;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGPPSLVCLLSATVFSLLGGSSAFSLSHRLKGRFQDRRNIRPNIIILVLTDDQDELGS 60  
 Db 1 MGPPSLVCLLSATVFSLLGGSSAFSLSHRLKGRFQDRRNIRPNIIILVLTDDQDELGS 60  
 QY 61 MQVMNKTTRIMEQGGTHFINAFVTPMCCPSRSSILTGKYVHNHTYTNNECSPSQWA 120  
 Db 61 MQVMNKTTRIMEQGGTHFINAFVTPMCCPSRSSILTGKYVHNHTYTNNECSPSQWA 120  
 QY 121 QHESRTFVYLINSTGYRTAFPGKYLYNEVNGSVPPGKWEVGLLKNRFRNYTLCRNGVK 180  
 Db 121 QHESRTFVYLINSTGYRTAFPGKYLYNEVNGSVPPGKWEVGLLKNRFRNYTLCRNGVK 180  
 QY 181 EXHGSYDYSKDYLTDLITNDSVSFFRTSKMTYPHRPLVLMVISHAAPHGPDSPAPQYSLRFP 240  
 Db 181 EXHGSYDYSKDYLTDLITNDSVSFFRTSKMTYPHRPLVLMVISHAAPHGPDSPAPQYSLRFP 240  
 QY 241 NASOHITPSYAPNDKHMIRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300  
 Db 241 NASOHITPSYAPNDKHMIRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300  
 QY 301 LVETGELDNITYIVYTADHGYHIGFGLVKGKSMPEYFDIRVPFVYRGPNVEAGCLNPHIV 360  
 Db 301 LVETGELDNITYIVYTADHGYHIGFGLVKGKSMPEYFDIRVPFVYRGPNVEAGCLNPHIV 360  
 QY 361 LNIDLAPTILDIAGLDIPADMDGKSILKLLDTERPVRNPHLKKQVRVWDSFLVERGKLL 420  
 Db 361 LNIDLAPTILDIAGLDIPADMDGKSILKLLDTERPVRNPHLKKQVRVWDSFLVERGKLL 420  
 QY 421 HKRDNKVDQAEENFLPKYQVRKDLQCBAEYOTACEQLGQKQWQVEDATGKLLHKCKGP 480

Db 421 HKRDNDKVDQAEEENFLPKYQRYKDLQRAEYQTAECQLGQKWQCVEDATGKLKHKCKGP 480

QY 481 MRLGGSRALSNLVPKYGGGSEACTCDSDYKLSLAGRRKLLFKKKYKASYRBSIRSV 540

Db 481 MRLGGSRALSNLVPKYGGGSEACTCDSDYKLSLAGRRKLLFKKKYKASYRBSIRSV 540

QY 541 AIEVDGRVYHVLGDAQAOPRNLTKRHWPAGPEQDDKDGDFSGTGGLPDYSAANPIKVT 600

Db 541 AIEVDGRVYHVLGDAQAOPRNLTKRHWPAGPEQDDKDGDFSGTGGLPDYSAANPIKVT 600

QY 601 HRCYILENDTVQCDLIDLYKSLQAWKHKLHIDHEIETLQNKIKNLRVGRHLKXKRPBEC 660

Db 601 HRCYILENDTVQCDLIDLYKSLQAWKHKLHIDHEIETLQNKIKNLRVGRHLKXKRPBEC 660

QY 661 DCHKISYHTQHKGRGLKHGSSLHPPFRKGLQKDKVWLLREOKRKKLKLKRLQNNDTC 720

Db 661 DCHKISYHTQHKGRGLKHGSSLHPPFRKGLQKDKVWLLREOKRKKLKLKRLQNNDTC 720

QY 721 SMPGLTCTFTHDQHWOTAPFWTLGPPFCACTSANNTYWCMTINETHNLFCEFATGFE 780

Db 721 SMPGLTCTFTHDQHWOTAPFWTLGPPFCACTSANNTYWCMTINETHNLFCEFATGFE 780

QY 781 YEDLNTDQYOLMNAVNTLDRVLNQLHVQLMELRSCGYKQCNPRTRNMDLGLKDGGSYE 840

Db 781 YEDLNTDQYOLMNAVNTLDRVLNQLHVQLMELRSCGYKQCNPRTRNMDLGLKDGGSYE 840

QY 841 QYRQQRKRWPEMKRPPSSKSLGQWEGWEG 870

Db 841 QYRQQRKRWPEMKRPPSSKSLGQWEGWEG 870

RESULT 4

AAAM79215

ID AAAM79215 standard; Protein; 870 AA.

AC AAAM79215;

XX 06-NOV-2001 (first entry)

XX Human protein SEQ ID NO 1877.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;

XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

XX tissue growth factor; immunomodulatory; cancer; leukaemia;

XX nervous system disorder; arthritis; inflammation.

OS Homo sapiens.

OS WO200157190-A2.

PN 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US04098.

XX 03-FEB-2000; 2000US-0496914.

PR 27-APR-2000; 2000US-0560875.

PR 20-JUN-2000; 2000US-0598075.

PR 19-JUL-2000; 2000US-0620325.

PR 01-SEP-2000; 2000US-0654936.

PR 15-SEP-2000; 2000US-0663561.

PR 20-OCT-2000; 2000US-0693325.

PR 30-NOV-2000; 2000US-0728422.

XX (HYSE-) HYSEQ INC.

PA Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;

PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX WPI; 2001-476283/51.

DR N-PSDB; AAK52348.

XX Nucleic acids encoding polypeptides with cytokine-like activities,

PT useful in diagnosis and gene therapy -

XX Claim 20; Page 4262-4264; 622lpp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation.

CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the sequence listing were missing at the time of publication.

XX

SQ Sequence 870 AA;

Query Match 99.9%; Score 4746; DB 22; Length 870;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGPPSLVLCILSATVFSLLGGSSAFLSHRLKGRFORDRRNIRPNILVLTDDQDVELGS 60

Db 1 MGPPSLVLCILSATVFSLLGGSSAFLSHRLKGRFORDRRNIRPNILVLTDDQDVELGS 60

QY 61 MQVMNKTTRIMEQGGTHFINAFVTTMCCPSRSSILTKYVHNHNTYTNNENCSFSPWA 120

Db 61 MQVMNKTTRIMEQGAHFINAFVTTMCCPSRSSILTKYVHNHNTYTNNENCSFSPWA 120

QY 121 QHESRTFAVYLNSTGYRTAFFGKYLINEYNGSVYPPGKWKVGLLKNSRFNYTLCRNGVK 180

Db 121 QHESRTFAVYLNSTGYRTAFFGKYLINEYNGSVYPPGKWKVGLLKNSRFNYTLCRNGVK 180

QY 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYRHPVLMVISHAAPHGSDAPQYSLRFP 240

Db 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYRHPVLMVISHAAPHGSDAPQYSLRFP 240

QY 241 NASQHTPSYNYAPNPKHMYRTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300

Db 241 NASQHTPSYNYAPNPKHMYRTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300

QY 301 LVETGELDNITYIVYTADHGVHIGQGLVKGKSNPYEFDIRVPEYVGGNVAGCLNPHIV 360

Db 301 LVETGELDNITYIVYTADHGVHIGQGLVKGKSNPYEFDIRVPEYVGGNVAGCLNPHIV 360

QY 361 LNIDLAPTILDIAGLDIPADMCKSILKLLDTERPVNRFHLKKMVRWDRSFLVERGKLL 420

Db 361 LNIDLAPTILDIAGLDIPADMCKSILKLLDTERPVNRFHLKKMVRWDRSFLVERGKLL 420

QY 421 HKRDNDKVDQAEEENFLPKYQRYKDLQRAEYQTAECQLGQKWQCVEDATGKLKHKCKGP 480

Db 421 HKRDNDKVDQAEEENFLPKYQRYKDLQRAEYQTAECQLGQKWQCVEDATGKLKHKCKGP 480

QY 481 MRLGGSRALSNLVPKYGGGSEACTCDSDYKLSLAGRRKLLFKKKYKASYRBSIRSV 540

Db 481 MRLGGSRALSNLVPKYGGGSEACTCDSDYKLSLAGRRKLLFKKKYKASYRBSIRSV 540

QY 541 AIEVDGRVYHVLGDAQAOPRNLTKRHWPAGPEQDDKDGDFSGTGGLPDYSAANPIKVT 600

Db 541 AIEVDGRVYHVLGDAQAOPRNLTKRHWPAGPEQDDKDGDFSGTGGLPDYSAANPIKVT 600

QY 601 HRCYILENDTVQCDLIDLYKSLQAWKHKLHIDHEIETLQNKIKNLRVGRHLKXKRPBEC 660

Db 601 HRCYILENDTVQCDLIDLYKSLQAWKHKLHIDHEIETLQNKIKNLRVGRHLKXKRPBEC 660

QY 661 DCHKISYHTQHKGRGLKHGSSLHPPFRKGLQKDKVWLLREOKRKKLKLKRLQNNDTC 720

Db 661 DCHKISYHTQHKGRGLKHGSSLHPPFRKGLQKDKVWLLREOKRKKLKLKRLQNNDTC 720

QY 721 SMPGLTCTFDHONQWOTAPFWTLGPFCACTSANNTYWCMTINETHNFCEPATGFL 780  
 Db |||||  
 QY 721 SMPGLTCTFDHONQWOTAPFWTLGPFCACTSANNTYWCMTINETHNFCEPATGFL 780  
 Db |||||  
 QY 781 YFDLNTDPYQLMNAVNTLDRVLNQLHVQLMELSCSKYKQCNPRTRNMGLGKDGGSYE 840  
 Db |||||  
 QY 841 QYRQPRKWPCKPSSKSLGQLWEGWEG 870  
 Db |||||  
 QY 841 QYRQPRKWPCKPSSKSLGQLWEGWEG 870  
 Db |||||

RESULT 5  
 AAE01440  
 ID AAE01440 standard; Protein; 870 AA.  
 XX AC  
 XX AAE01440;  
 XX DT 17-JUL-2001 (first entry)  
 XX DE Human gene 5 encoded secreted protein HE9QN39, SEQ ID NO:95.  
 XX KW Human; secreted protein; proliferative disorder; cancer; tumour;  
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;  
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;  
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;  
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;  
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;  
 KW cardiovascular disorder; angiogenic disorder; kidney disorder;  
 KW gastrointestinal disorder; pregnancy-related disorder;  
 KW endocrine disorder; infection; wound healing; vulnerability;  
 KW cell culture; chemotaxis; food additive; gene therapy;  
 KW binding partner identification.  
 XX OS Homo sapiens.  
 XX FH Key  
 FT Peptide 1..24  
 FT /label= Signal\_peptide  
 FT Protein 25..870  
 FT /note= "Mature human secreted protein"  
 XX WO200134626-A1.  
 XX PD 17-MAY-2001.  
 XX 01-NOV-2000; 2000WO-US30045.  
 XX 05-NOV-1999; 99US-0163581.  
 XX 30-JUN-2000; 2000US-0215133.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX Ruben SM, Komatsoulis GA, Moore PA, Birse CE, Ni J;  
 WPI; 2001-308778/32.  
 DR N-PSDB; AAD05304.  
 XX New nucleic acid molecules encoding 28 human secreted proteins for  
 PT diagnosing, preventing, treating or ameliorating medical conditions and  
 PT used as food additives or preservatives -  
 XX Claim 11; Page 486-489; 562pp; English.  
 XX AAD05300-AAD05379 represent cDNAs corresponding to 28 human secreted  
 CC protein genes, and AAE01436-AAE01513 represent the proteins or variants.  
 CC AAE01514-AAE01544 represent human secreted protein fragments or variants.  
 CC The genes and their secreted proteins are useful for preventing,  
 CC treating or ameliorating medical conditions, e.g., by protein or gene  
 CC therapy. Pathological conditions can be diagnosed by determining the  
 CC amount of the new protein in a sample or by determining the presence of  
 CC mutations in the new genes. Specific uses are described for each of the

CC 28 genes, based on the tissues in which they are most highly expressed,  
 CC and include developing products for the diagnosis or treatment of  
 CC proliferative disorders, cancer, tumours, foetal and developmental  
 CC abnormalities, haematopoietic disorders, diseases of the immune system,  
 CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,  
 CC allergies, neurological disorders (e.g., Alzheimer's disease,  
 CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,  
 CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,  
 CC cardiovascular disorders, angiogenic disorders, kidney disorders,  
 CC gastrointestinal disorders, pregnancy-related disorders, endocrine  
 CC disorders, and infections. The proteins can also be used to aid wound  
 CC healing and epithelial cell proliferation, to prevent skin aging due to  
 CC sunburn, to maintain organs before transplantation, for supporting cell  
 CC culture of primary tissues, to regenerate tissues, to identify their  
 CC cognate ligands or binding partners, and in chemotaxis, and can be used  
 CC as a food additive or preservative to modify storage properties.  
 CC Antibodies specific for a protein of the invention can be used in  
 CC alleviating symptoms associated with the disorders mentioned above, and  
 CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked  
 CC immunosorbent assay (ELISA). The present sequence represents a human  
 CC secreted protein of the invention.  
 XX SQ Sequence 870 AA;

Query Match 99.9%; Score 4746; DB 22; Length 870;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGPPSLVLCLLSATVFSLLGGSSAFSLSHRLKGRFQDRNRIRNPILVLTDDQDVELGS 60  
 Db |||||  
 QY 1 MGPPSLVLCLLSATVFSLLGGSSAFSLSHRLKGRFQDRNRIRNPILVLTDDQDVELGS 60  
 Db |||||  
 QY 61 MQVMKTRIRIMEOGGTHFINAFVTPMCCPSRSILITQKYVHNHTYNNENCSPPSHQA 120  
 Db |||||  
 QY 61 MQVMKTRIRIMEOGGAHFINAFVTPMCCPSRSILITQKYVHNHTYNNENCSPPSHQA 120  
 Db |||||  
 QY 121 QHESRTFAVILNSTGYRTAFFGKYLYNEVNGSVVPPGKWEVGLLKNSSFYNTLCRNVK 180  
 Db |||||  
 QY 121 QHESRTFAVILNSTGYRTAFFGKYLYNEVNGSVVPPGKWEVGLLKNSSFYNTLCRNVK 180  
 Db |||||  
 QY 181 EKHSYDYSKDYLTDLITNDSVSFFRTSKMTYPHRPVLMI SHAAPHGPEDSAPQYSLFP 240  
 Db |||||  
 QY 181 EKHSYDYSKDYLTDLITNDSVSFFRTSKMTYPHRPVLMI SHAAPHGPEDSAPQYSLFP 240  
 Db |||||  
 QY 241 NASQHTPSYVYAPNDKHWIMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNNM 300  
 Db |||||  
 QY 241 NASQHTPSYVYAPNDKHWIMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNNM 300  
 Db |||||  
 QY 301 LVETGELDNTYIVYTADHGYHIGFGLVKGXSMPEYFDIRVPFVYRGNVEAGCLNPHIV 360  
 Db |||||  
 QY 301 LVETGELDNTYIVYTADHGYHIGFGLVKGXSMPEYFDIRVPFVYRGNVEAGCLNPHIV 360  
 Db |||||  
 QY 361 LNIDLAPTILDIAGLDIPADMDGKSILKLLDTERPVNRPHLKKOORVWRDSFLVERGKLL 420  
 Db |||||  
 QY 361 LNIDLAPTILDIAGLDIPADMDGKSILKLLDTERPVNRPHLKKOORVWRDSFLVERGKLL 420  
 Db |||||  
 QY 421 HKRDNKVDYDAOBENFLPKYQVYKDLQRAEYOTACEOLGQKWCQVEDATGKLKHKCKGP 480  
 Db |||||  
 QY 421 HKRDNKVDYDAOBENFLPKYQVYKDLQRAEYOTACEOLGQKWCQVEDATGKLKHKCKGP 480  
 Db |||||  
 QY 481 MRLGGSRALSNLVPKYQGGSEACTCDSGDYKLSLAGRRKLFKKYKASYRSRSIRSV 540  
 Db |||||  
 QY 481 MRLGGSRALSNLVPKYQGGSEACTCDSGDYKLSLAGRRKLFKKYKASYRSRSIRSV 540  
 Db |||||  
 QY 541 ALEVDGRVYHVLGDAAOPRNLTKRHWPAGAPEDQDDKGGDFSGTGLPDYSAANPKVT 600  
 Db |||||  
 QY 541 ALEVDGRVYHVLGDAAOPRNLTKRHWPAGAPEDQDDKGGDFSGTGLPDYSAANPKVT 600  
 Db |||||  
 QY 601 HRCYILENDTVQCDDLDLYKSLQAKMDKHLHIDHIEITLQNKIKLREVRGHLKKRPEEC 660  
 Db |||||  
 QY 601 HRCYILENDTVQCDDLDLYKSLQAKMDKHLHIDHIEITLQNKIKLREVRGHLKKRPEEC 660  
 Db |||||  
 QY 661 DCHKISYHTQHGRKLRGSSSLHPFRKGLQEKQVWLLREQKRRKKLKLRLQNNDTC 720

Db 661 DCHKLSYHTQHKGRKLRGSSLPFRKGLQEKDKVLLREKQKQKRLKRLKRLQNNDC 720  
 Qy 721 SMPGLTCTFHDHNOHWOTAPFWTLGPPFCACTSANNNTYWCMRTINETHNPLFCEPATGFLE 780  
 Db 721 SMPGLTCTFHDHNOHWOTAPFWTLGPPFCACTSANNNTYWCMRTINETHNPLFCEPATGFLE 780  
 Qy 781 YFDLNTDPPVLMNANVTLDRLVNLQHLVQLMELRSCKGYKQCNPRTRMDLGLKDGGSYE 840  
 Db 781 YFDLNTDPPVLMNANVTLDRLVNLQHLVQLMELRSCKGYKQCNPRTRMDLGLKDGGSYE 840  
 Qy 841 QYRQFORRWPEMKRPPSSKSLQQLWEGWEG 870  
 Db 841 QYRQFORRWPEMKRPPSSKSLQQLWEGWEG 870

## RESULT 6

ABG71836

ID ABG71836 standard; Protein; 870 AA.

XX ABG71836;

AC ABG71836;

DT 20-JAN-2003 (first entry)

DE human SULF2 protein.

KW Human; SULF2; glucosamine-6-sulphatase; cancer; ischaemia; enzyme;  
 KW tumour; angiogenesis; coronary; carotid; arterial occlusive disease;  
 KW peripheral arterial disease; atherosclerosis; myointimal hyperplasia;  
 KW thromboangitis obliterans; thrombotic disorder; vasculitis;  
 KW heart attack; myocardial infarction; vascular death; inflammation;  
 KW rheumatoid arthritis; asthma; adult respiratory distress syndrome;  
 KW sarcoidosis; hypersensitivity pneumonitis; multiple sclerosis;  
 KW allograft rejection; lymphoma; thrombosis; sulphatase.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..22

FT /note= "Signal peptide"

FT Protein 25..870

FT /note= "Mature protein"

FT Cleavage-site 409..410

FT /note= "furan/PACE protease cleavage site"

FT Cleavage-site 423..424

FT /note= "furan/PACE protease cleavage site"

FT Cleavage-site 538..539

FT /note= "furan/PACE protease cleavage site"

FT Cleavage-site 565..566

FT /note= "furan/PACE protease cleavage site"

XX WO200259327-A2.

XX 01-AUG-2002.

XX 26-DEC-2001; 2001WO-US49793.

XX 27-DEC-2000; 2000US-258577P.

XX 09-FEB-2001; 2001US-267831P.

XX (REGC ) UNIV CALIFORNIA.

XX Rosen SD, Hemmerich S, Tomita M, Palmeri D;

XX WPI; 2002-636541/68.

XX N-PSDB; ABS56553.

XX New sulfatase polypeptides having glucosamine-6-sulfatase activity,  
 XX useful in screening, discovery and preparation of diagnostic and  
 XX therapeutic agents for treating cancer, ischemic conditions,  
 XX inflammation, or thrombosis

XX Example 1; Fig2A; 293pp; English.

PS

XX This invention relates to the DNA and protein sequences of a novel  
 CC polypeptide having glucosamine-6-sulphatase activity. The sulphatases of  
 CC are useful in screening, discovery and preparation of diagnostic and  
 CC therapeutic agents for treating cancer, ischaemic conditions,  
 CC inflammation, or thrombosis. The nucleic acids are useful in preparing  
 CC the sulfatase polypeptides, identifying the expression of genes in a  
 CC biological specimen, or generating transgenic non-human animals or  
 CC site-specific gene modification in cell lines. The host cells are  
 CC useful in replicating and/or expressing the polynucleotides or nucleic  
 CC acids. The agents are useful in treating the disorders cited above by  
 CC reducing tumour growth, inflammation, and thrombosis, or increasing  
 CC angiogenesis, e.g. by treating coronary, carotid, or arterial occlusive  
 CC disease, peripheral arterial disease, atherosclerosis, myointimal  
 CC hyperplasia, thromboangitis obliterans, thrombotic disorders,  
 CC vasculitis; or preventing ischemic conditions, heart attack (myocardial  
 CC infarction), or other vascular death. The sulphatases and/or agents are  
 CC also useful in treating rheumatoid arthritis, asthma, adult respiratory  
 CC distress syndrome, sarcoidosis, hypersensitivity pneumonitis, multiple  
 CC sclerosis, allograft rejection, and spread of lymphomas to cutaneous  
 CC sites. The present sequence represents the human SULF2 protein of the  
 CC invention.  
 XX  
 SQ Sequence 870 AA;

Query Match 99.9%; Score 4746; DB 23; Length 870;

Best Local Similarity 99.9%; Pred. NO. 0;

Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGPPSLVCLLSATVFSLLGSSAFSLSHRLKGRFQDRNRNIRPNILVLTDDQDVELGS 60

Db 1 MGPPSLVCLLSATVFSLLGSSAFSLSHRLKGRFQDRNRNIRPNILVLTDDQDVELGS 60

Qy 61 MQVNNKTRIRIMEQGQAHFINAFVTTMCCPSSSILTKYVHNHNTYTNENCSSPSWA 120

Db 61 MQVNNKTRIRIMEQGQAHFINAFVTTMCCPSSSILTKYVHNHNTYTNENCSSPSWA 120

Qy 121 QHESRTPAVYLNSTGYRTAFFGKYLNEYNGSVYPPGKWEVGLLKNSRFYNYTLCRNGVK 180

Db 121 QHESRTPAVYLNSTGYRTAFFGKYLNEYNGSVYPPGKWEVGLLKNSRFYNYTLCRNGVK 180

Qy 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKKMYPHRVLWVISHAAPHGSDAPQYSLRFP 240

Db 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKKMYPHRVLWVISHAAPHGSDAPQYSLRFP 240

Qy 241 NASOHLTPSYNYAPNDKGMWIMRYTGMKPIHMEFTNMLQKLEOTLMSVDDSMETIYNN 300

Db 241 NASOHLTPSYNYAPNDKGMWIMRYTGMKPIHMEFTNMLQKLEOTLMSVDDSMETIYNN 300

Qy 301 LVETGELDNTYIIVYTADHGYHIQFGLVKGSMPYEFDIRVPYVVRGPNVVEAGCLNPHIV 360

Db 301 LVETGELDNTYIIVYTADHGYHIQFGLVKGSMPYEFDIRVPYVVRGPNVVEAGCLNPHIV 360

Qy 361 LNTDLAPTILDIAGLOIPADMDGKSLKLLDTERPVNRPHLKKQMRVWRDSEFLVERGKLL 420

Db 361 LNTDLAPTILDIAGLOIPADMDGKSLKLLDTERPVNRPHLKKQMRVWRDSEFLVERGKLL 420

Qy 421 HKRDNKVDQAQEEFNLPKYORVKDLCORAEYQACQLGQKQOCVEDATGKLLHKCKGP 480

Db 421 HKRDNKVDQAQEEFNLPKYORVKDLCORAEYQACQLGQKQOCVEDATGKLLHKCKGP 480

Qy 481 MRLGGSRALSNLVPKYGGSGSEACTCDSGDYKLSLAGRRKKLFKKYKASYVRSRSIRSV 540

Db 481 MRLGGSRALSNLVPKYGGSGSEACTCDSGDYKLSLAGRRKKLFKKYKASYVRSRSIRSV 540

Qy 541 AIEVDGRVHVGLGDAAPRNLTKRHWPGAPEDDDKDGDFSGTGGLPDYSAANPIKVT 600

Db 541 AIEVDGRVHVGLGDAAPRNLTKRHWPGAPEDDDKDGDFSGTGGLPDYSAANPIKVT 600

Qy 601 HRCVILENDTVQCDLILYKSLQAKDKHLHDIETLQNKILNREVRGHLKKCKPEEC 660

Db 601 HRCVILENDTVQCDLILYKSLQAKDKHLHDIETLQNKILNREVRGHLKKCKPEEC 660

QY 661 DCHKISYHTQHKRLKXRGSSLPFRKGLQEKDQVWLLRQKRRKKLRLKRLQNNNTC 720  
 D 661 DCHKISYHTQHKRLKXRGSSLPFRKGLQEKDQVWLLRQKRRKKLRLKRLQNNNTC 720  
 QY 721 SMPGLTCFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNPLFCERATGLE 780  
 D 721 SMPGLTCFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNPLFCERATGLE 780  
 QY 781 YFOLNTDPYQLMNAVNTLDRDVLNQLHVQLMEILRSCKYKQCNPRTRNMDLGLKDGGSYE 840  
 D 781 YFOLNTDPYQLMNAVNTLDRDVLNQLHVQLMEILRSCKYKQCNPRTRNMDLGLKDGGSYE 840  
 QY 841 QYRQFQRKWPENKRPSSKSLGOLWEGWEG 870  
 D 841 QYRQFQRKWPENKRPSSKSLGOLWEGWEG 870

RESULT 7  
 ID ABG63903  
 AC ABG63903;  
 XX 27-AUG-2002 (first entry)  
 DT Human albumin fusion protein #578.  
 XX  
 DE Albumin fusion protein; therapeutic protein X; human albumin; HA;  
 XX human serum albumin; HSA; cancer; reproductive disorder;  
 KW digestive disorder; immune disorder; endocrine disorder;  
 KW haematopoietic disorder; neural disorder; connective disorder;  
 KW cytotatic; antiinfertility; antiinflammatory; antiulcer;  
 KW immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;  
 KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;  
 KW osteopathic; antiarthritic.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 OS WO20017137-A1.  
 PN  
 XX 18-OCT-2001.  
 PD  
 XX 12-APR-2001; 2001WO-US11988.  
 PF  
 XX 12-APR-2000; 2000US-229358P.  
 PR 25-APR-2000; 2000US-199384P.  
 PR 21-DEC-2000; 2000US-256931P.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 XX Rosen CA, Haseltine WA;  
 PI WPI; 2002-010886/01.  
 XX  
 XX New fusion protein for treating disease e.g. diabetes comprises an  
 PT albumin fused to a therapeutic protein -  
 XX  
 XX Claim 1; Page 892-894; 2102pp; English.  
 PS  
 XX The present invention relates to albumin fusion proteins comprising a  
 CC therapeutic protein X and human albumin (HA, also known as human serum  
 CC albumin, HSA). The proteins are useful for treating a disease or  
 CC disorder that may be modulated by therapeutic protein X. The albumin  
 CC extends the shelf-life of protein X, and may increase its biological  
 CC in vitro/in vivo activity. The protein is useful for treating and  
 CC diagnosing disorders such as cancer, reproductive disorders, digestive  
 CC disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders  
 CC (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders  
 CC (e.g. diabetes), haematopoietic disorders, neural disorders  
 CC (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,  
 CC encephalomyelitis, meningitis, schizophrenia), and connective disorders  
 CC (e.g. osteoporosis, arthritis). ABG63326-ABG65518 represent albumin

CC fusion proteins of the invention.  
 XX  
 SQ Sequence 870 AA;  
 Query Match 99.9%; Score 4746; DB 23, Length 870;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MGPPSLVCLLSATVFSLLGSSAFSLSHRLKGRFORRRNIRPNILVLTDODDELGS 60  
 D 1 MGPPSLVCLLSATVFSLLGSSAFSLSHRLKGRFORRRNIRPNILVLTDODDELGS 60  
 QY 61 MQVMNTRIMEQGGTHFINAFVTTMCCPSRSSILTKYVHNNTYNNENCSPPSQA 120  
 D 61 MQVMNTRIMEQGGAHFINAFVTTMCCPSRSSILTKYVHNNTYNNENCSPPSQA 120  
 QY 121 QHESRTFAVLNSTGYRTAFGKYLNEVNGSVYPPGKWEVGLLKNRSFYNTLCRNGVK 180  
 D 121 QHESRTFAVLNSTGYRTAFGKYLNEVNGSVYPPGKWEVGLLKNRSFYNTLCRNGVK 180  
 QY 181 EXHGSYSDYLTDLITNDSVSFFRTSKMYPHRPVLMVI SHAAHPGDESAPOYSRLFP 240  
 D 181 EXHGSYSDYLTDLITNDSVSFFRTSKMYPHRPVLMVI SHAAHPGDESAPOYSRLFP 240  
 QY 241 NASOHLTPSYNYPAPNDKHMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300  
 D 241 NASOHLTPSYNYPAPNDKHMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300  
 QY 301 LVETGELDNITYIVYADHGYHIGFGLVKGKSMPEYEDIRVPFYVRGPNVEAGCLNPHIV 360  
 D 301 LVETGELDNITYIVYADHGYHIGFGLVKGKSMPEYEDIRVPFYVRGPNVEAGCLNPHIV 360  
 QY 361 LNIIDLAPTILDTAGLDIPADMDGKSLILKLDTERPVNRFHKKKRWVRDSEFLVERGKLL 420  
 D 361 LNIIDLAPTILDTAGLDIPADMDGKSLILKLDTERPVNRFHKKKRWVRDSEFLVERGKLL 420  
 QY 421 HKRDNDKVDQAQENFLPKYQRVKOLCORAEYQTAQCEQLGQKQWOCVEDATGKLHKCKGP 480  
 D 421 HKRDNDKVDQAQENFLPKYQRVKOLCORAEYQTAQCEQLGQKQWOCVEDATGKLHKCKGP 480  
 QY 481 MRLGGSRALSNLVPKYQGGSEACTCDSGDYKLSLAGRRKKLFKKYKASYRSTRSV 540  
 D 481 MRLGGSRALSNLVPKYQGGSEACTCDSGDYKLSLAGRRKKLFKKYKASYRSTRSV 540  
 QY 541 AIEVDGRVYHVLGDAAPRNLTKEHWPAGPDODDQDGGDFSGTGGLPDYSAANPKVT 600  
 D 541 AIEVDGRVYHVLGDAAPRNLTKEHWPAGPDODDQDGGDFSGTGGLPDYSAANPKVT 600  
 QY 601 HRCYILENDTVQCDLDLYKSLQAWKDKHLHDHETLQNKIKNLREVRGHLKKRPEEC 660  
 D 601 HRCYILENDTVQCDLDLYKSLQAWKDKHLHDHETLQNKIKNLREVRGHLKKRPEEC 660  
 QY 661 DCHKISYHTQHKRLKXRGSSLPFRKGLQEKDQVWLLRQKRRKKLRLKRLQNNNTC 720  
 D 661 DCHKISYHTQHKRLKXRGSSLPFRKGLQEKDQVWLLRQKRRKKLRLKRLQNNNTC 720  
 QY 721 SMPGLTCFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNPLFCERATGLE 780  
 D 721 SMPGLTCFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNPLFCERATGLE 780  
 QY 781 YFOLNTDPYQLMNAVNTLDRDVLNQLHVQLMEILRSCKYKQCNPRTRNMDLGLKDGGSYE 840  
 D 781 YFOLNTDPYQLMNAVNTLDRDVLNQLHVQLMEILRSCKYKQCNPRTRNMDLGLKDGGSYE 840  
 QY 841 QYRQFQRKWPENKRPSSKSLGOLWEGWEG 870  
 D 841 QYRQFQRKWPENKRPSSKSLGOLWEGWEG 870

RESULT 8  
 ID AAE01471  
 XX AAE01471 standard; Protein; 870 AA.

AC AAE01471;  
 DT 17-JUL-2001 (first entry)  
 XX Human gene 5 encoded secreted protein HE9QN39, SEQ ID NO:127.  
 XX Human; secreted protein; proliferative disorder; cancer; tumour;  
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;  
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;  
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;  
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;  
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;  
 KW cardiovascular disorder; angiodysplasia; kidney disorder;  
 KW gastrointestinal disorder; pregnancy-related disorder;  
 KW endocrine disorder; infection; wound healing; vulnery;  
 KW cell culture; chemotaxis; food additive; gene therapy;  
 KW binding partner identification.  
 XX  
 OS Homo sapiens.  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..24  
 FT /label= Signal\_peptide  
 FT Protein 25..870  
 FT /note= "Mature human secreted protein"  
 FT Misc-difference 537  
 FT /label= Unknown  
 FT /note= "Encoded by ART"  
 XX  
 PN WO200134626-A1.  
 XX  
 PD 17-MAY-2001.  
 XX  
 PF 01-NOV-2000; 2000WO-US30045.  
 XX  
 XX 05-NOV-1999; 99US-0163581.  
 PR 30-JUN-2000; 2000US-0215133.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 XX Ruben SM, Komatsoulis GA, Moore PA, Birse CE, Ni J;  
 PI WP2; 2001-308778/32.  
 DR N-PSDB; AAD01471.  
 XX  
 PT New nucleic acid molecules encoding 28 human secreted proteins for  
 PT diagnosing, preventing, treating or ameliorating medical conditions and  
 PT used as food additives or preservatives -  
 XX  
 PS Claim 11; Page 507-510; 562pp; English.  
 CC AAD05300-AAD05379 represent cDNAs corresponding to 28 human secreted  
 CC protein genes, and AAE01436-AAE01513 represent the proteins they encode.  
 CC AAE01514-AAE01544 represent human secreted protein fragments or variants.  
 CC The genes and their secreted proteins are useful for preventing,  
 CC treating or ameliorating medical conditions, e.g., by protein or gene  
 CC therapy. Pathological conditions can be diagnosed by determining the  
 CC amount of the new protein in a sample or by determining the presence of  
 CC mutations in the new genes. Specific uses are described for each of the  
 CC 28 genes, based on the tissues in which they are most highly expressed,  
 CC and include developing products for the diagnosis or treatment of  
 CC proliferative disorders, cancer, tumours, foetal and developmental  
 CC abnormalities, haematopoietic disorders, diseases of the immune system,  
 CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,  
 CC allergies, neurological disorders (e.g., Alzheimer's disease,  
 CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,  
 CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,  
 CC cardiovascular disorders, angiodysplasia, kidney disorders,  
 CC gastrointestinal disorders, pregnancy-related disorders, endocrine  
 CC disorders, and infections. The proteins can also be used to aid wound  
 CC healing and epithelial cell proliferation, to prevent skin aging due to  
 CC sunburn, to maintain organs before transplantation, for supporting cell  
 CC culture of primary tissues, to regenerate tissues, to identify their

CC cognate ligands or binding partners, and in chemotaxis, and can be used  
 CC as a food additive or preservative to modify storage properties.  
 CC Antibodies specific for a protein of the invention can be used in  
 CC alleviating symptoms associated with the disorders mentioned above, and  
 CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked  
 CC immunosorbent assay (ELISA). The present sequence represents a human  
 CC secreted protein of the invention.  
 XX  
 SQ Sequence 870 AA;

Query Match 99.8%; Score 4742; DB 22; Length 870;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 868; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MGPPSLVCLLSATVPSLLGGSSAPLSHRLKGRFDRNRINPNILVLTDQDVELGS 60  
 Db 1 MGPPSLVCLLSATVPSLLGGSSAPLSHRLKGRFDRNRINPNILVLTDQDVELGS 60  
 Qy 61 MQVMNKTTRIMEGGTHFINAFVTTMCCPSRSSILTKGVVHNHTYTNENCSSPSWQA 120  
 Db 61 MQVMNKTTRIMEGGGAHFINAFVTTMCCPSRSSILTKGVVHNHTYTNENCSSPSWQA 120  
 Qy 121 QHESRTFAVYLNSTGYRTAFFGKYLYNEVNGSVYPPGKWEVGLLKNSRFNYTLCENGVK 180  
 Db 121 QHESRTFAVYLNSTGYRTAFFGKYLYNEVNGSVYPPGKWEVGLLKNSRFNYTLCENGVK 180  
 Qy 181 EKIGSDYSKDYLTDLITNDSVSFFRTSKOMYPHRPVLMI SHAAPHGPDSDAPQSRLEP 240  
 Db 181 EKIGSDYSKDYLTDLITNDSVSFFRTSKOMYPHRPVLMI SHAAPHGPDSDAPQSRLEP 240  
 Qy 241 NASQHITPSVNYAPNPKWIMRYTGPMPKPIHMEFTNMLQRLQTLMSVDDSMETIYNN 300  
 Db 241 NASQHITPSVNYAPNPKWIMRYTGPMPKPIHMEFTNMLQRLQTLMSVDDSMETIYNN 300  
 Qy 301 LVETGELDNTYIVYTADHGVIHQFGLVKGKSNPYEFDIRVPFYVRGPNVEAGLNPHIV 360  
 Db 301 LVETGELDNTYIVYTADHGVIHQFGLVKGKSNPYEFDIRVPFYVRGPNVEAGLNPHIV 360  
 Qy 361 LNTDLAPTLDIAGLDIPADMDGKSLKLLDTERPVNRFHLKKQRRVWRDVSFLVERGKLL 420  
 Db 361 LNTDLAPTLDIAGLDIPADMDGKSLKLLDTERPVNRFHLKKQRRVWRDVSFLVERGKLL 420  
 Qy 421 HKRDNDKVAQAEENFLPKYQVRVNDLCQRAYOTACBOLGQWQCVEDATGKLLHKCKGP 480  
 Db 421 HKRDNDKVAQAEENFLPKYQVRVNDLCQRAYOTACBOLGQWQCVEDATGKLLHKCKGP 480  
 Qy 481 MRLGGSRALSNLVPKYVGQSGSEACTCDSDYKLSLAGRRKULFKKKYKASYVRSIRSV 540  
 Db 481 MRLGGSRALSNLVPKYVGQSGSEACTCDSDYKLSLAGRRKULFKKKYKASYVRSIRSV 540  
 Qy 541 AIEVDGRVHVGLGDAAPRNLTKRHWPGAPEDDDKGGDFSGTGGLPDYSAANPIKYT 600  
 Db 541 AIEVDGRVHVGLGDAAPRNLTKRHWPGAPEDDDKGGDFSGTGGLPDYSAANPIKYT 600  
 Qy 601 HRCYILENDTVQCDLDLYKSLQAKDHLHIHIEITLQNKIKNLRVGRHLKKRPEEC 660  
 Db 601 HRCYILENDTVQCDLDLYKSLQAKDHLHIHIEITLQNKIKNLRVGRHLKKRPEEC 660  
 Qy 661 DCHKISYHTQHKRLKHGSSLLHPFKGLOBKQKWLIREQKQKULKLLRLQNNDTFC 720  
 Db 661 DCHKISYHTQHKRLKHGSSLLHPFKGLOBKQKWLIREQKQKULKLLRLQNNDTFC 720  
 Qy 721 SMPGLTCFTHDNHWTAPFWTLGPPCACTSANNTYWCMTINETHNLFCEFAFGFUE 780  
 Db 721 SMPGLTCFTHDNHWTAPFWTLGPPCACTSANNTYWCMTINETHNLFCEFAFGFUE 780  
 Qy 781 YFDLNTDPQLMANVNTLDRVLNQLHVLQMLRSCYKQCNPRTRNMDLKGXGGSYE 840  
 Db 781 YFDLNTDPQLMANVNTLDRVLNQLHVLQMLRSCYKQCNPRTRNMDLKGXGGSYE 840  
 Qy 841 QYRQFORRWKPEMKRPPSSKSLQOLMEGWEG 870  
 Db 841 QYRQFORRWKPEMKRPPSSKSLQOLMEGWEG 870

RESULT 9  
 ABG63902  
 ID ABG63902 standard; Protein; 870 AA.  
 XX  
 AC ABG63902;  
 XX  
 DT 27-AUG-2002 (first entry)  
 XX  
 DE Human albumin fusion protein #577.  
 XX  
 KW Albumin fusion protein; therapeutic protein X; human albumin; HA;  
 KW human serum albumin; HSA; cancer; reproductive disorder;  
 KW digestive disorder; immune disorder; endocrine disorder;  
 KW haematopoietic disorder; neural disorder; connective disorder;  
 KW cytostatic; antiinfertility; antiinflammatory; antiulcer;  
 KW immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;  
 KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;  
 KW osteopathic; antiarthritic.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO2001177137-A1.  
 XX  
 PD 18-OCT-2001.  
 XX  
 PF 12-APR-2001; 2001WO-US11988.  
 XX  
 PR 12-APR-2000; 2000US-229358P.  
 PR 25-APR-2000; 2000US-199384P.  
 PR 21-DEC-2000; 2000US-256931P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Haseltine WA;  
 XX  
 DR WPI; 2002-010886/01.  
 XX  
 XX  
 PT New fusion protein for treating disease e.g. diabetes comprises an  
 PT albumin fused to a therapeutic protein -  
 XX  
 PS Claim 1; Page 889-891; 2102pp; English.  
 XX  
 CC The present invention relates to albumin fusion proteins comprising a  
 CC therapeutic protein X and human albumin (HA), also known as human serum  
 CC albumin, HSA). The proteins are useful for treating a disease or  
 CC disorder that may be modulated by therapeutic protein X. The albumin  
 CC extends the shelf-life of protein X, and may increase its biological  
 CC in vitro/in vivo activity. The protein is useful for treating and  
 CC diagnosing disorders such as cancer, reproductive disorders, digestive  
 CC disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders  
 CC (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders  
 CC (e.g. diabetes), haematopoietic disorders, neural disorders  
 CC (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,  
 CC encephalomyelitis, meningitis, schizophrenia), and connective disorders  
 CC (e.g. osteoporosis, arthritis). ABG63326-ABG65518 represent albumin  
 CC fusion proteins of the invention.  
 XX  
 XX Sequence 870 AA;  
 SQ

Query Match 99.8%; Score 4742; DB 23; Length 870;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 868; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MGPPSLVCLLSATVPSLGGSSAFSLSHHLLKGRFORRRNIRPNILVLTDQDVELGS 60  
 Db 1 MGPPSLVCLLSATVPSLGGSSAFSLSHHLLKGRFORRRNIRPNILVLTDQDVELGS 60  
 Qy 61 MQVNNKTRRIMEQGGPHFNAFVTPMCCPSRSSILTKYVHNHNTYNNENCSPPSQA 120  
 Db 61 MQVNNKTRRIMEQGGPHFNAFVTPMCCPSRSSILTKYVHNHNTYNNENCSPPSQA 120

Qy 121 QHSRTFAVYLNSTGYRTAFPGKYLNEYNGSYVPPGKEMWGLLKNSRFNYTLCRNGVK 180  
 Db 121 QHSRTFAVYLNSTGYRTAFPGKYLNEYNGSYVPPGKEMWGLLKNSRFNYTLCRNGVK 180  
 Qy 181 EKHSYDYSKYLTDLITNDSSVSFRTSKMYPHRPVLMVISHAAPHGPDSPAYSLFF 240  
 Db 181 EKHSYDYSKYLTDLITNDSSVSFRTSKMYPHRPVLMVISHAAPHGPDSPAYSLFF 240  
 Qy 241 NASOHITPSYNYAPNPKHIMRYTGPMKPIHMEFTNNLQKRLQTLMSVDDSMETIYNN 300  
 Db 241 NASOHITPSYNYAPNPKHIMRYTGPMKPIHMEFTNNLQKRLQTLMSVDDSMETIYNN 300  
 Qy 301 LVETGELDNTYIVVTADHYHIGFGLVKGSMPEYFDIRVPFYVRGPNVEAGCLNPHIV 360  
 Db 301 LVETGELDNTYIVVTADHYHIGFGLVKGSMPEYFDIRVPFYVRGPNVEAGCLNPHIV 360  
 Qy 361 LNIIDLAPTILDIAGLDIPADMDGKSIILKLDTERPVNRFLHKKOMRVRRDSFLVERKLL 420  
 Db 361 LNIIDLAPTILDIAGLDIPADMDGKSIILKLDTERPVNRFLHKKOMRVRRDSFLVERKLL 420  
 Qy 421 HKRDNDKVDAQEENFLPKYQRYVDLCQAEYQTACEQLGQKWQCVEDATGKLHKCKGP 480  
 Db 421 HKRDNDKVDAQEENFLPKYQRYVDLCQAEYQTACEQLGQKWQCVEDATGKLHKCKGP 480  
 Qy 481 MRLGGSRALSNLVPKYQGSGSEACTCDSDYKLSLAGRRKKLFKKYKASYVRSRSIRSV 540  
 Db 481 MRLGGSRALSNLVPKYQGSGSEACTCDSDYKLSLAGRRKKLFKKYKASYVRSRSIRSV 540  
 Qy 541 AIEVDGRVYHVLGDAAPRNLTKRHPGAPEDQDDKDGDFSGTGLPDYSAANPIKVT 600  
 Db 541 AIEVDGRVYHVLGDAAPRNLTKRHPGAPEDQDDKDGDFSGTGLPDYSAANPIKVT 600  
 Qy 601 HRCYILENDTVQCDLDLYKSLQAWKHKLHDHIEITLQNKIKVLREVRGHLKKRPEEC 660  
 Db 601 HRCYILENDTVQCDLDLYKSLQAWKHKLHDHIEITLQNKIKVLREVRGHLKKRPEEC 660  
 Qy 661 DCHKISYHTQHKRLKRGSSSLHPRKGLQEKDKVLLREOKKKLKLKRLQNNDTG 720  
 Db 661 DCHKISYHTQHKRLKRGSSSLHPRKGLQEKDKVLLREOKKKLKLKRLQNNDTG 720  
 Qy 721 SMPGLTCFTHDNQHWQTAPFWTLGPFCACTSANNTYVCMRTINETNFIPLCFEATGFE 780  
 Db 721 SMPGLTCFTHDNQHWQTAPFWTLGPFCACTSANNTYVCMRTINETNFIPLCFEATGFE 780  
 Qy 781 YFDLNTDPVQLMNAVNTLDRDVLNQLHVQLMELSSCKGYKOCNPRTRNMDLGLKDGGSYE 840  
 Db 781 YFDLNTDPVQLMNAVNTLDRDVLNQLHVQLMELSSCKGYKOCNPRTRNMDLGLKDGGSYE 840  
 Qy 841 QYRQFQRKRWPEMKRPPSSKSLGQLWEGWEG 870  
 Db 841 QYRQFQRKRWPEMKRPPSSKSLGQLWEGWEG 870

RESULT 10  
 AB990378  
 ID AB990378 standard; Protein; 870 AA.  
 XX  
 AC AB990378;  
 XX  
 DT 24-MAY-2002 (first entry)  
 XX  
 DE Human polypeptide SEQ ID NO 2754.  
 XX  
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;  
 KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;  
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; human; secreted protein.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200190304-A2.



XX PD 29-NOV-2001. XX  
XX PF 18-MAY-2001; 2001WO-US16450. XX  
XX PF 19-MAY-2000; 2000US-205515P. XX  
XX PA (HUMA-) HUMAN GENOME SCI INC. XX  
XX PI Birse CE, Rosen CA; XX  
XX DR WPI; 2002-122018/16. XX  
XX DR N-PSDB; ABL90787. XX  
XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and  
PT prevention of neural, immune system, muscular, reproductive,  
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative  
PT disorders -  
XX Claim 11; SEQ ID NO 2754; 2081pp + Sequence Listing; English.  
XX The invention relates to novel genes (ABL89449-ABL90853) and proteins  
CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy. The genes are  
CC isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;  
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC and parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX SQ Sequence 870 AA;  
Query Match 99.8%; Score 4742; DB 23; Length 870;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 868; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 MGPPSLVLCCLLSATVPSLLGSSAFUSHRLKGRFORRRNIRPNILVLTDQDVELGS 60  
DB 1 MGPPSLVLCCLLSATVPSLLGSSAFUSHRLKGRFORRRNIRPNILVLTDQDVELGS 60  
QY 61 MOVNKTTRIMEGGTHFTNAFVTPMCCPSSILTGKYVHNHTYNNENCSFSAQA 120  
DB 61 MOVNKTTRIMEGGHAFFNAFVTPMCCPSSILTGKYVHNHTYNNENCSFSAQA 120  
QY 121 QHESRTFAVYLNSTGYRTAFFGKYLNEYGSYVPPGKWEVGLLKNRSFYNTLCRNGVK 180  
DB 121 QHESRTFAVYLNSTGYRTAFFGKYLNEYGSYVPPGKWEVGLLKNRSFYNTLCRNGVK 180  
QY 181 EXHGSYSDYLDLTITNDVSFFRTSKYKPHRPVLWVISHAAPHGSDSAPQYSLRFP 240  
DB 181 EXHGSYSDYLDLTITNDVSFFRTSKYKPHRPVLWVISHAAPHGSDSAPQYSLRFP 240  
QY 241 NASQHTTPSYNAPNDKXWIMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDSMETIYNM 300  
DB 241 NASQHTTPSYNAPNDKXWIMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDSMETIYNM 300  
QY 301 LVETGELDNITYIYVADHGHVIGQFGLVKGKSNPYEFDIRVPPYVGRPNVEAGCLNPHIV 360  
DB 301 LVETGELDNITYIYVADHGHVIGQFGLVKGKSNPYEFDIRVPPYVGRPNVEAGCLNPHIV 360  
QY 361 LNTDLAPTILDIAGLDIPADMDGKSLKLLDTERPVRNPHLKKRVRWDSFLVERGKLL 420  
DB 361 LNTDLAPTILDIAGLDIPADMDGKSLKLLDTERPVRNPHLKKRVRWDSFLVERGKLL 420

QY 421 HKRDNDKVDQAQENFLPKYQRYKDLQRAEYQTAQELGQKMQCQVEDATGKLKHKCKGP 480  
DB 421 HKRDNDKVDQAQENFLPKYQRYKDLQRAEYQTAQELGQKMQCQVEDATGKLKHKCKGP 480  
QY 481 MRLGGSRLSNLVPKYGGSEACTCDSDGYKSLAGRRKKLKKYKASYVRSRSIRSV 540  
DB 481 MRLGGSRLSNLVPKYGGSEACTCDSDGYKSLAGRRKKLKKYKASYVRSRSIRSV 540  
QY 541 AIEVDGRVYHVGLDAAQPRNLTQKHWPGAPEDQDDKGGDFSGTGGGLPDYSAANPIKVT 600  
DB 541 AIEVDGRVYHVGLDAAQPRNLTQKHWPGAPEDQDDKGGDFSGTGGGLPDYSAANPIKVT 600  
QY 601 HRCYILENDTVQCDLDLYKSLQAMKDKHLHDHEIETLQNKIKNLREVRGHLKKRPEEC 660  
DB 601 HRCYILENDTVQCDLDLYKSLQAMKDKHLHDHEIETLQNKIKNLREVRGHLKKRPEEC 660  
QY 661 DCHKISYHTQHKGRGLKRGSSSLHPRKGLQEKQVWLLREQKKKKLKLKRLQNDTC 720  
DB 661 DCHKISYHTQHKGRGLKRGSSSLHPRKGLQEKQVWLLREQKKKKLKLKRLQNDTC 720  
QY 721 SMPGLTCFTHDQHWQTAPFWTLGPFCACTSANNTYWCRTINETHFLFCFATGFLE 780  
DB 721 SMPGLTCFTHDQHWQTAPFWTLGPFCACTSANNTYWCRTINETHFLFCFATGFLE 780  
QY 781 YPDLTNDPYQLANVNTLDRDVLNQLHVMELRSCKGYKQCNPRTRNMDLGLKGGSYE 840  
DB 781 YPDLTNDPYQLANVNTLDRDVLNQLHVMELRSCKGYKQCNPRTRNMDLGLKGGSYE 840  
QY 841 QYRQFORRWPMKRPSSKSLGQLWEGWEG 870  
DB 841 QYRQFORRWPMKRPSSKSLGQLWEGWEG 870  
RESULT 11  
AAAY66648  
ID AAAY66648 standard; protein; 867 AA.  
XX AC AAAY66648;  
XX DT 05-APR-2000 (first entry)  
XX DE Membrane-bound protein PRO1120.  
XX KM Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;  
XX KM pharmaceutical; receptor immunoadhesin; gene mapping.  
XX OS Homo sapiens.  
XX PN WO9963088-A2.  
XX PD 09-DEC-1999.  
XX PF 02-JUN-1999; 99WO-US12252.  
XX PR 02-JUN-1998; 98US-0087607.  
XX PR 02-JUN-1998; 98US-0087609.  
XX PR 02-JUN-1998; 98US-0087759.  
XX PR 03-JUN-1998; 98US-0087827.  
XX PR 04-JUN-1998; 98US-0088021.  
XX PR 04-JUN-1998; 98US-0088025.  
XX PR 04-JUN-1998; 98US-0088028.  
XX PR 04-JUN-1998; 98US-0088029.  
XX PR 04-JUN-1998; 98US-0088030.  
XX PR 04-JUN-1998; 98US-0088033.  
XX PR 04-JUN-1998; 98US-0088326.  
XX PR 05-JUN-1998; 98US-0088167.  
XX PR 05-JUN-1998; 98US-0088202.  
XX PR 05-JUN-1998; 98US-0088217.  
XX PR 09-JUN-1998; 98US-0088655.  
XX PR 10-JUN-1998; 98US-0088722.  
XX PR 10-JUN-1998; 98US-0088730.

PR 10-JUN-1998; 98US-0088734.  
PR 10-JUN-1998; 98US-0088738.  
PR 10-JUN-1998; 98US-0088740.  
PR 10-JUN-1998; 98US-0088741.  
PR 10-JUN-1998; 98US-0088742.  
PR 10-JUN-1998; 98US-0088810.  
PR 10-JUN-1998; 98US-0088811.  
PR 10-JUN-1998; 98US-0088824.  
PR 10-JUN-1998; 98US-0088825.  
PR 10-JUN-1998; 98US-0088826.  
PR 11-JUN-1998; 98US-0088858.  
PR 11-JUN-1998; 98US-0088861.  
PR 11-JUN-1998; 98US-0088863.  
PR 11-JUN-1998; 98US-0088876.  
PR 12-JUN-1998; 98US-0089090.  
PR 12-JUN-1998; 98US-0089105.  
PR 16-JUN-1998; 98US-0089440.  
PR 16-JUN-1998; 98US-0089512.  
PR 16-JUN-1998; 98US-0089514.  
PR 17-JUN-1998; 98US-0089532.  
PR 17-JUN-1998; 98US-0089538.  
PR 17-JUN-1998; 98US-0089598.  
PR 17-JUN-1998; 98US-0089599.  
PR 17-JUN-1998; 98US-0089600.  
PR 17-JUN-1998; 98US-0089653.  
PR 18-JUN-1998; 98US-0089801.  
PR 18-JUN-1998; 98US-0089807.  
PR 18-JUN-1998; 98US-0089908.  
PR 19-JUN-1998; 98US-0089947.  
PR 19-JUN-1998; 98US-0089948.  
PR 19-JUN-1998; 98US-0089952.  
PR 22-JUN-1998; 98US-0090246.  
PR 22-JUN-1998; 98US-0090252.  
PR 23-JUN-1998; 98US-0090254.  
PR 23-JUN-1998; 98US-0090349.  
PR 23-JUN-1998; 98US-0090355.  
PR 24-JUN-1998; 98US-0090429.  
PR 24-JUN-1998; 98US-0090431.  
PR 24-JUN-1998; 98US-0090435.  
PR 24-JUN-1998; 98US-0090444.  
PR 24-JUN-1998; 98US-0090445.  
PR 24-JUN-1998; 98US-0090461.  
PR 24-JUN-1998; 98US-0090472.  
PR 24-JUN-1998; 98US-0090535.  
PR 24-JUN-1998; 98US-0090538.  
PR 24-JUN-1998; 98US-0090540.  
PR 25-JUN-1998; 98US-0090557.  
PR 25-JUN-1998; 98US-0090676.  
PR 25-JUN-1998; 98US-0090678.  
PR 25-JUN-1998; 98US-0090688.  
PR 25-JUN-1998; 98US-0090690.  
PR 25-JUN-1998; 98US-0090691.  
PR 25-JUN-1998; 98US-0090694.  
PR 25-JUN-1998; 98US-0090695.  
PR 25-JUN-1998; 98US-0090696.  
PR 26-JUN-1998; 98US-0090862.  
PR 26-JUN-1998; 98US-0090863.  
PR 01-JUL-1998; 98US-0091358.  
PR 01-JUL-1998; 98US-0091360.  
PR 01-JUL-1998; 98US-0091544.  
PR 02-JUL-1998; 98US-0091478.  
PR 02-JUL-1998; 98US-0091486.  
PR 02-JUL-1998; 98US-0091519.  
PR 02-JUL-1998; 98US-0091626.  
PR 02-JUL-1998; 98US-0091628.  
PR 02-JUL-1998; 98US-0091633.  
PR 02-JUL-1998; 98US-0091634.  
PR 02-JUL-1998; 98US-0091646.  
PR 02-JUL-1998; 98US-0091673.  
PR 07-JUL-1998; 98US-0091978.  
PR 09-JUL-1998; 98US-0091982.  
PR 10-JUL-1998; 98US-0092182.  
PR 20-JUL-1998; 98US-0092472.  
PR 20-JUL-1998; 98US-0093339.

PR 30-JUL-1998; 98US-0094651.  
PR 04-AUG-1998; 98US-0095282.  
PR 04-AUG-1998; 98US-0095285.  
PR 04-AUG-1998; 98US-0095301.  
PR 04-AUG-1998; 98US-0095302.  
PR 04-AUG-1998; 98US-0095308.  
PR 04-AUG-1998; 98US-0095321.  
PR 04-AUG-1998; 98US-0095325.  
PR 10-AUG-1998; 98US-0095916.  
PR 10-AUG-1998; 98US-0095929.  
PR 10-AUG-1998; 98US-0096012.  
PR 11-AUG-1998; 98US-0096143.  
PR 11-AUG-1998; 98US-0096146.  
PR 12-AUG-1998; 98US-0096329.  
PR 17-AUG-1998; 98US-0096757.  
PR 17-AUG-1998; 98US-0096766.  
PR 17-AUG-1998; 98US-0096768.  
PR 17-AUG-1998; 98US-0096773.  
PR 17-AUG-1998; 98US-0096791.  
PR 17-AUG-1998; 98US-0096867.  
PR 17-AUG-1998; 98US-0096891.  
PR 17-AUG-1998; 98US-0096894.  
PR 17-AUG-1998; 98US-0096895.  
PR 17-AUG-1998; 98US-0096897.  
PR 18-AUG-1998; 98US-0096949.  
PR 18-AUG-1998; 98US-0096950.  
PR 18-AUG-1998; 98US-0096959.  
PR 18-AUG-1998; 98US-0096960.  
PR 18-AUG-1998; 98US-0097022.  
PR 19-AUG-1998; 98US-0097141.  
PR 20-AUG-1998; 98US-0097218.  
PR 24-AUG-1998; 98US-0097661.  
PR 26-AUG-1998; 98US-0097951.  
PR 26-AUG-1998; 98US-0097952.  
PR 26-AUG-1998; 98US-0097954.  
PR 26-AUG-1998; 98US-0097955.  
PR 26-AUG-1998; 98US-0097971.  
PR 26-AUG-1998; 98US-0097974.  
PR 26-AUG-1998; 98US-0097978.  
PR 26-AUG-1998; 98US-0097979.  
PR 26-AUG-1998; 98US-0097986.  
PR 26-AUG-1998; 98US-0098014.  
PR 31-AUG-1998; 98US-0098525.  
PR 16-SEP-1998; 98US-0100634.  
PR 12-JAN-1999; 99US-0115565.  
XX  
PA (GETH ) GENENTECH INC.  
PI Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;  
PI Wood WI, Yuan J;  
XX  
XX WPI; 2000-072883/06.  
DR N-PSDB; AAZ64969.  
XX  
XX Membrane-bound proteins and related nucleotide sequences -  
XX  
XX claim 12; Fig 47; 822pp; English.  
XX  
XX The invention provides membrane-bound PRO polypeptides and  
CC polynucleotides encoding them. The PRO sequences of the invention were  
CC identified based on extracellular domain homology screening. The PRO  
CC sequences have homology with proteins including LDL receptors, TIE  
CC ligands and various enzymes. The membrane-bound proteins and receptor  
CC molecules are useful as pharmaceutical and diagnostic agents. Receptor  
CC immunoadhesins, for instance, can be used as therapeutic agents to block  
CC receptor-ligand interactions. The membrane-bound proteins can also be  
CC employed for screening of potential peptide or small molecule inhibitors  
CC of the relevant receptor/ligand interaction. The PRO encoding sequences  
CC are useful as hybridization probes, in chromosome and gene mapping and in  
CC the generation of antisense RNA and DNA. PRO nucleic acid sequences  
CC will also be useful for the preparation of PRO polypeptides, especially  
CC by recombinant techniques.

SQ	Sequence	867 AA;
	Query Match	99.3%; Score 4719.5; DB 21; Length 867;
	Best Local Similarity	99.5%; Pred. No. 0;
	Matches	866; Conservative 0; Mismatches 1; Indels 3; Gaps 1;
QY	1	MGPPSLVLCLLSATVFLSGSSAFSLHRLKGRFORDRNRIPNIIILVLTDDQVLEGS 60
DB	1	MGPPSLVLCLLSATVFLSGSSAFSLHRLKGRFORDRNRIPNIIILVLTDDQVLEGS 60
QY	61	MOVNKKTRRIMEQGGTHFINAFVTTMCCPSSRSILTGKYVHNHTYTNNECSPSQWA 120
DB	61	MOVNKKTRRIMEQGGTHFINAFVTTMCCPSSRSILTGKYVHNHTYTNNECSPSQWA 120
QY	121	QHSRTFAVLNLTGTYRTAFGKYLNEYSGYVPPGKEMVGLLKNSRFYNTLCRNGVK 180
DB	121	QHSRTFAVLNLTGTYRTAFGKYLNEYSGYVPPGKEMVGLLKNSRFYNTLCRNGVK 180
QY	181	EKGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAHPGDESDAPQYSLRFP 240
DB	181	EKGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAHPGDESDAPQYSLRFP 240
QY	241	NASQHTPSNYAPNPKWIMYTGPMKPIHMEFTNMLQKELQTLMSVDSMETIYNN 300
DB	241	NASQHTPSNYAPNPKWIMYTGPMKPIHMEFTNMLQKELQTLMSVDSMETIYNN 300
QY	301	LVETGELDNTYIVYTADHGVIHGGFLVKGKSNPYEFDIRVPPYVPGNVEAGCLNPHIV 360
DB	301	LVETGELDNTYIVYTADHGVIHGGFLVKGKSNPYEFDIRVPPYVPGNVEAGCLNPHIV 360
QY	361	LNIDLAPTLIDIAGLDIPADMDGKSILKLDTERPVNRPHLKKOMVRWDSFLVERGKLL 420
DB	361	LNIDLAPTLIDIAGLDIPADMDGKSILKLDTERPVNRPHLKKOMVRWDSFLVERGKLL 420
QY	421	HKRDNKDVAQENFLPKYQKVDLCQAEYQACQLOKQWQCVEDATGKLLHKCKGP 480
DB	421	HKRDNKDVAQENFLPKYQKVDLCQAEYQACQLOKQWQCVEDATGKLLHKCKGP 480
QY	481	MRLGSRALSNLVPKYGGQSEACTCDSDYKLSLAGRBKLPKKYKASYRSTRSIV 540
DB	481	MRLGSRALSNLVPKYGGQSEACTCDSDYKLSLAGRBKLPKKYKASYRSTRSIV 540
QY	541	AIEDVGRVYHVLGDAAPRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKYT 600
DB	541	AIEDVGRVYHVLGDAAPRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKYT 600
QY	601	HRCYLENDTVQCDLDLYKSIQAKDKHLHIHIEITLQNKIKNLEVRGHLKKRPEEC 660
DB	601	HRCYLENDTVQCDLDLYKSIQAKDKHLHIHIEITLQNKIKNLEVRGHLKKRPEEC 660
QY	661	DCHKISYHTQKGRKLRHSGSLHPPFRKGLQEKQKVLRLREQKPKKLRKLLKRLQNNDT 720
DB	661	DCHKISYHTQKGRKLRHSGSLHPPFRKGLQEKQKVLRLREQKPKKLRKLLKRLQNNDT 720
QY	721	SMPGLTCTFDHNOHWOTAPFWTLGPPCACSANNNTYWCRTINETHNLFCEPATGELE 780
DB	721	SMPGLTCTFDHNOHWOTAPFWTLGPPCACSANNNTYWCRTINETHNLFCEPATGELE 780
QY	781	YFDLNTDTPYQLANAVNTLDRLVNLQHLVQLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840
DB	781	YFDLNTDTPYQLANAVNTLDRLVNLQHLVQLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840
QY	841	QYRQFQRKPKWPKRPSKSLQGLWEGBG 870
DB	838	QYRQFQRKPKWPKRPSKSLQGLWEGBG 867

RESULT 12

ID AAU12181 standard; Protein; 867 AA.

XX AAU12181;

XX

DT	24-OCT-2001	(first entry)
XX	Human PRO1120 polypeptide sequence.	
XX	Human secretory and transmembrane; PRO; mammalian; cancer; lung;	
KW	breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;	
KW	cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;	
KW	adipocyte; A-peptide; factor VIIA; gene therapy.	
XX	Homo sapiens.	
OS	WO200140466-A2.	
XX	07-JUN-2001.	
XX	01-DEC-2000; 2000WO-US32678.	
XX	01-DEC-1999; 99WO-US28301.	
PR	01-DEC-1999; 99WO-US28634.	
PR	02-DEC-1999; 99WO-US28551.	
PR	02-DEC-1999; 99WO-US28564.	
PR	02-DEC-1999; 99WO-US28565.	
PR	09-DEC-1999; 99US-0170262.	
PR	16-DEC-1999; 99WO-US30095.	
PR	20-DEC-1999; 99WO-US30911.	
PR	20-DEC-1999; 99WO-US30999.	
PR	30-DEC-1999; 99WO-US31243.	
PR	06-JAN-2000; 2000WO-US00277.	
PR	06-JAN-2000; 2000WO-US00376.	
PR	11-FEB-2000; 2000WO-US03565.	
PR	18-FEB-2000; 2000WO-US04341.	
PR	18-FEB-2000; 2000WO-US04342.	
PR	22-FEB-2000; 2000WO-US04414.	
PR	24-FEB-2000; 2000WO-US04914.	
PR	24-FEB-2000; 2000WO-US05004.	
PR	01-MAR-2000; 2000WO-US05601.	
PR	20-MAR-2000; 2000WO-US07377.	
PR	21-MAR-2000; 2000WO-US07532.	
PR	30-MAR-2000; 2000WO-US08439.	
PR	17-MAY-2000; 2000WO-US13705.	
PR	22-MAY-2000; 2000WO-US14042.	
PR	30-MAY-2000; 2000WO-US14941.	
PR	02-JUN-2000; 2000WO-US15264.	
PR	10-NOV-2000; 2000WO-US30873.	
XX	(GETH ) GENENTECH INC.	
PA	Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;	
XX	Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;	
PI	Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;	
XX	WPI; 2001-408281/43.	
DR	N-PSDB; AAS21253.	
XX	Isolated, secretory and transmembrane PRO polypeptide used to detect	
PT	other PRO polypeptides, link bioactive molecules to cells expressing	
PT	PRO polypeptides, and detect the presence of mammalian tumours e.g.	
PT	lung, breast, prostate, cervical	
XX	Claim 12; Fig 20; 813pp; English.	
XX	AAU12172-AAU12446 represent novel human secretory and transmembrane	
CC	PRO polypeptides. The PRO polypeptides are useful to detect other	
CC	PRO polypeptides, to link bioactive molecules to cells expressing	
CC	PRO polypeptides, to modulate biological activities of cells expressing	
CC	PRO polypeptides, and to detect the presence of mammalian lung, colon,	
CC	breast, prostate, rectal, cervical or liver tumours by comparing PRO	
CC	polypeptide expression in a cell sample to that in a control sample.	
CC	Some of the 275 sequences are also useful to stimulate the release of	
CC	tumour necrosis factor-alpha (TNF-alpha) from human blood, the	
CC	proliferation or differentiation of chondrocytes, the proliferation or	
CC	gene expression in pericyte cells, the release of proteoglycans from	
CC	cartilage, the proliferation of inner ear utricular supporting cells or	

CC of T-lymphocytes, the release of a cytokine from peripheral blood  
 CC monocytes (PMCs), or the proliferation of endothelial cells. Some of  
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by  
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide  
 CC to factor VIIA. The PRO polypeptides can be used in assays to identify  
 CC molecules involved in binding interactions. The polynucleotides encoding  
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,  
 CC transgenic or knock out animals and can be used in gene therapy.  
 XX  
 SQ Sequence 867 AA;  
 Query Match 99.3%; Score 4719.5; DB 22; Length 867;  
 Best Local Similarity 99.5%; Pred. No. 0;  
 Matches 866; Conservative 0; Mismatches 1; Indels 3; Gaps 1;  
 QY 1 MGPPSLVCLLSATVFLSGSSAFSLSHRLKGRFQDRDNIRPNILVLTDDQDELGS 60  
 DB 1 MGPPSLVCLLSATVFLSGSSAFSLSHRLKGRFQDRDNIRPNILVLTDDQDELGS 60  
 QY 61 MOVNKKTRRIMEQGTTHFINAVTTPMCCPSRSSILTKYVHNHTYTNNECSSPSMQA 120  
 DB 61 MOVNKKTRRIMEQGAHFINAVTTPMCCPSRSSILTKYVHNHTYTNNECSSPSMQA 120  
 QY 121 QHESRTFAVILNSTGYRTAFPGKYLNEYSGYVPPGKEMVGLLKNSRFNYTLCRNGVK 180  
 DB 121 QHESRTFAVILNSTGYRTAFPGKYLNEYSGYVPPGKEMVGLLKNSRFNYTLCRNGVK 180  
 QY 181 EKHGSDYSKDYLDLITNDVSFSPRTSKQMYPHRPVLMVISHAAPHGPDSPAFQSRLLFP 240  
 DB 181 EKHGSDYSKDYLDLITNDVSFSPRTSKQMYPHRPVLMVISHAAPHGPDSPAFQSRLLFP 240  
 QY 241 NASQHITPSYVADPKHIMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNM 300  
 DB 241 NASQHITPSYVADPKHIMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNM 300  
 QY 301 LVETGELDNTYIVYADHGYHIGQGLVKGKSMPEYEDIRVPFYVGPVNEAGCLNPHIV 360  
 DB 301 LVETGELDNTYIVYADHGYHIGQGLVKGKSMPEYEDIRVPFYVGPVNEAGCLNPHIV 360  
 QY 361 LNIDLAPTLIDLAGLIDPADMDGKSILKLDTERPVRNPHLKKQVRWRDSFLVGRKLL 420  
 DB 361 LNIDLAPTLIDLAGLIDPADMDGKSILKLDTERPVRNPHLKKQVRWRDSFLVGRKLL 420  
 QY 421 HKRNDKVDQAEENFLPKYQVKDLQRAEYQTAQCEQLGQKMQCQVEDATGKLLHKCKGP 480  
 DB 421 HKRNDKVDQAEENFLPKYQVKDLQRAEYQTAQCEQLGQKMQCQVEDATGKLLHKCKGP 480  
 QY 481 MRLGSSRALSNLVPKYQGSSEACTDSGDYKLSLAGRRKLLFKKTKASYVRSRSIRSV 540  
 DB 481 MRLGSSRALSNLVPKYQGSSEACTDSGDYKLSLAGRRKLLFKKTKASYVRSRSIRSV 540  
 QY 541 ALEVDRVYHVLGDAQAPRLTKRWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600  
 DB 541 ALEVDRVYHVLGDAQAPRLTKRWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600  
 QY 601 HRCYILENDTVQCDLDLYKSLQAKWKHLHDHIEITLQNKIKWLREVRGHLKKQKPEEC 660  
 DB 601 HRCYILENDTVQCDLDLYKSLQAKWKHLHDHIEITLQNKIKWLREVRGHLKKQKPEEC 660  
 QY 661 DCHKLSYTHQHKRLKRGSSLHPFRKGLQEKDKVLLREOKRKKLKLKRLQNNDDTC 720  
 DB 661 DCHKLSYTHQHKRLKRGSSLHPFRKGLQEKDKVLLREOKRKKLKLKRLQNNDDTC 720  
 QY 721 SMPGLTCTFDHNDHQWQTAPFTWLGPFCACTSANNTYWCMTINETHNLFCEFAFGLE 780  
 DB 721 SMPGLTCTFDHNDHQWQTAPFTWLGPFCACTSANNTYWCMTINETHNLFCEFAFGLE 780  
 QY 781 YFDLNTDYPQLMNAVNTLDRVNLQHLVQLMELRSCGYKQCNPRTRNMDLGLKDGSSYE 840  
 DB 781 YFDLNTDYPQLMNAVNTLDRVNLQHLVQLMELRSCGYKQCNPRTRNMDLGLKDGSSYE 840  
 QY 841 QYRQFQRKWPMPKRPSSKSLQQLMEGWEG 870

DB 838 QYRQFQRKWPMPKRPSSKSLQQLMEGWEG 867  
 RESULT 13  
 AAB65171  
 ID AAB65171 standard; Protein; 867 AA.  
 XX  
 AC AAB65171;  
 XX  
 DT 02-APR-2001 (first entry)  
 XX  
 DE Human PRO1120 (UNQ559) protein sequence SEQ ID NO:84.  
 XX  
 KW Human; secreted and transmembrane protein; PRO; cytotstatic;  
 KW cell death; cancer; chromosomal mapping; gene mapping; tissue typing;  
 KW diagnostic assay.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO2000073454-A1.  
 XX  
 PD 07-DEC-2000.  
 XX  
 PF 30-MAR-2000; 2000WO-US08439.  
 XX  
 PR 02-JUN-1999; 99WO-US12252.  
 PR 23-JUN-1999; 99US-0141037.  
 PR 07-JUL-1999; 99US-0143048.  
 PR 20-JUL-1999; 99US-0144758.  
 PR 26-JUL-1999; 99US-0145698.  
 PR 28-JUL-1999; 99US-0146222.  
 PR 17-AUG-1999; 99US-0149396.  
 PR 15-SEP-1999; 99MO-US21090.  
 PR 15-SEP-1999; 99MO-US21547.  
 PR 08-OCT-1999; 99US-0158663.  
 PR 30-NOV-1999; 99MO-US28313.  
 PR 01-DEC-1999; 99MO-US28301.  
 PR 16-DEC-1999; 99MO-US30095.  
 PR 20-DEC-1999; 99MO-US30911.  
 PR 05-JAN-2000; 2000WO-US00219.  
 PR 06-JAN-2000; 2000WO-US00376.  
 PR 11-FEB-2000; 2000WO-US03565.  
 PR 18-FEB-2000; 2000WO-US04341.  
 PR 22-FEB-2000; 2000WO-US04414.  
 PR 24-FEB-2000; 2000WO-US04914.  
 PR 24-FEB-2000; 2000WO-US05004.  
 PR 02-MAR-2000; 2000WO-US05841.  
 PR 15-MAR-2000; 2000WO-US06884.  
 PR 20-MAR-2000; 2000WO-US07377.  
 XX  
 (GETH ) GENENTECH INC.  
 XX  
 AShtenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;  
 PI Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;  
 PI Grimaldi CJ, Gurney AJ, Kljavin IJ, Napier MA, Pan J, Paoni NF;  
 PI Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;  
 PI Zhang Z;  
 XX  
 WPI; 2001-032160/04.  
 DR N-PSDB; AAF44115.  
 XX  
 PRO polynucleotides used to produce polypeptides used to target  
 PT bioactive molecules such as toxins, radiolabels or antibodies, to  
 PT specific cells, to cause targeted cell death -  
 XX  
 Claim 12; Fig 47; 935pp; English.  
 XX  
 The present invention describes human secreted and transmembrane PRO  
 CC proteins. The PRO proteins have cytostatic activity. The PRO proteins  
 CC can be used for targeted delivery of bioactive molecules, such as  
 CC toxins, radiolabels or antibodies, that cause cell death. PRO nucleotide  
 CC sequences, and their fragments, can be used as hybridisation probes, in  
 CC chromosomal and gene mapping, and in the generation of anti-sense RNA

CC and DNA. They may also be used to produce transgenic animals which are  
 CC used to develop and screen therapeutically useful reagents. The PRO  
 CC nucleotide and protein sequence can be used for tissue typing and in  
 CC treating cancer. Anti-PRO antibodies can be used in diagnostic assays.  
 CC AAF44270 to AAF4470 represent PCR primers and hybridisation probes used  
 CC in the isolation of human PRO sequences. AAF44087 to AAF44269 and  
 CC AAB65154 to AAB65300 represent human PRO polynucleotide and protein  
 CC sequences given in the exemplification of the present invention.

XX Sequence 867 AA;

Query Match	99.3%;	Score	4719.5;	DB	22;	Length	867;
Best Local Similarity	99.5%;	Pred. No.	0;				
Matches	866;	Conservative	0;	Mismatches	1;	Indels	3;
						Gaps	1;

Qy 1 MGPPSLVCLLSATVPSLLGGSAFLSHRLKGRFORDRNIRPNILVLTDQDVELGS 60  
 Db 1 MGPPSLVCLLSATVPSLLGGSAFLSHRLKGRFORDRNIRPNILVLTDQDVELGS 60  
 Qy 61 MQVNMKTRIMEGGGTHFNAFVTTMCCPSRSSILTGKYYHNHTYNNENCSPPSWQA 120  
 Db 61 MQVNMKTRIMEGGGTHFNAFVTTMCCPSRSSILTGKYYHNHTYNNENCSPPSWQA 120  
 Qy 121 QHESRTFAVYLNSTGYRTAFYGGKLYNEVNGSVYPPGKWEVGLLKNRSFYNTLCRNGVK 180  
 Db 121 QHESRTFAVYLNSTGYRTAFYGGKLYNEVNGSVYPPGKWEVGLLKNRSFYNTLCRNGVK 180  
 Qy 181 EKHGSDYSKDYLTDLITNDVSFFRTSKXYPHRLVYISHAAPHGPDSPAYSRFLP 240  
 Db 181 EKHGSDYSKDYLTDLITNDVSFFRTSKXYPHRLVYISHAAPHGPDSPAYSRFLP 240  
 Qy 241 NASCHITPSNYAPNPKGMIWYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300  
 Db 241 NASCHITPSNYAPNPKGMIWYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300  
 Qy 301 LVETGELDNITYIYVYADHGVHIGQGLVKGKSNPYEFDIRVPPYVGRPNVEAGCLNPHIV 360  
 Db 301 LVETGELDNITYIYVYADHGVHIGQGLVKGKSNPYEFDIRVPPYVGRPNVEAGCLNPHIV 360  
 Qy 361 LNTDLAPTILDIAGLDIPADMDCKSLKLLDTERPVNREHKLKRVWRDSEFLVERGKLL 420  
 Db 361 LNTDLAPTILDIAGLDIPADMDCKSLKLLDTERPVNREHKLKRVWRDSEFLVERGKLL 420  
 Qy 421 HKRDNDKVDQAEENFLPKYQVXDLQCORAEYQACEQLGQKQCVEDATGKLKHKCKGP 480  
 Db 421 HKRDNDKVDQAEENFLPKYQVXDLQCORAEYQACEQLGQKQCVEDATGKLKHKCKGP 480  
 Qy 481 MRLGGSRALSNLVPKYGGGSEACTCDSDYKLSLAGREKLLFKKKYKASYRSHSISV 540  
 Db 481 MRLGGSRALSNLVPKYGGGSEACTCDSDYKLSLAGREKLLFKKKYKASYRSHSISV 540  
 Qy 541 A1EVDGRVYHVLGDAQAQPNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600  
 Db 541 A1EVDGRVYHVLGDAQAQPNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600  
 Qy 601 HRCVILENDTVQCDLDLYKSLQAWKHKLHIHIEITLQNKIKNIREVRGHKKRPEC 660  
 Db 601 HRCVILENDTVQCDLDLYKSLQAWKHKLHIHIEITLQNKIKNIREVRGHKKRPEC 660  
 Qy 661 DCHKISVHTQHKGRKLRHSGSLHPPFKGLQEKDKVWLLREQRKKLKLKRLQNNDT 720  
 Db 661 DCHKISVHTQHKGRKLRHSGSLHPPFKGLQEKDKVWLLREQRKKLKLKRLQNNDT 720  
 Qy 721 SMPGLTCTFDHNOHWQAPFWTILGPFCACTSANNTYWCMTINETHNPLFCEFATGPLE 780  
 Db 721 SMPGLTCTFDHNOHWQAPFWTILGPFCACTSANNTYWCMTINETHNPLFCEFATGPLE 780  
 Qy 781 YFDLNTDYPQLMNAVNTLDRLVNLQHLVOLMELRSCKYKQCNPRTRMDLGLKDGGSYE 840  
 Db 781 YFDLNTDYPQLMNAVNTLDRLVNLQHLVOLMELRSCKYKQCNPRTRMDLGLKDGGSYE 840  
 Qy 841 QYRQFQRKWKPEMKRSPSSKSLGLQWEGWEG 870  
 Db 841 QYRQFQRKWKPEMKRSPSSKSLGLQWEGWEG 870

Db	838	QYRQFQRKWKPEMKRSPSSKSLGLQWEGWEG	867
RESULT 14			
ABBS95459			
ID	ABBS95459	standard; Protein; 867 AA.	
XX	AC	ABB95459;	
XX	DT	19-JUL-2002 (first entry)	
XX	DE	Human angiogenesis related protein PRO120 SEQ ID NO: 74.	
XX	KW	Human; angiogenesis; PRO protein; cardiovascularisation; wound; cancer;	
XX	KW	atherosclerosis; cardiac hypertrophy; gene therapy; endothelial disorder;	
XX	KW	cardiant; cytostatic; antiangiogenic; hypotensive; vulnerary;	
XX	XX	antiarterioeclerotic.	
OS		Homo sapiens.	
XX	PN	WO200208284-A2.	
XX	PD	31-JAN-2002.	
XX	PF	09-JUL-2001; 2001WO-US21735.	
XX	PR	20-JUL-2000; 2000US-219556P.	
XX	PR	25-JUL-2000; 2000US-220624P.	
XX	PR	25-JUL-2000; 2000US-220664P.	
XX	PR	28-JUL-2000; 2000WO-US20710.	
XX	PR	02-AUG-2000; 2000US-222695P.	
XX	PR	17-AUG-2000; 2000US-0643657.	
XX	PR	23-AUG-2000; 2000WO-US23522.	
XX	PR	24-AUG-2000; 2000WO-US23328.	
XX	PR	07-SEP-2000; 2000US-230978P.	
XX	PR	15-SEP-2000; 2000US-000000P.	
XX	PR	18-SEP-2000; 2000US-0664610.	
XX	PR	18-SEP-2000; 2000US-0665350.	
XX	PR	24-OCT-2000; 2000US-242922P.	
XX	PR	08-NOV-2000; 2000US-0709238.	
XX	PR	08-NOV-2000; 2000WO-US30952.	
XX	PR	10-NOV-2000; 2000WO-US30873.	
XX	PR	20-DEC-2000; 2000US-0747259.	
XX	PR	20-DEC-2000; 2000WO-US34956.	
XX	PR	22-JAN-2001; 2000US-0767609.	
XX	PR	28-FEB-2001; 2001US-0796498.	
XX	PR	28-FEB-2001; 2001WO-US08520.	
XX	PR	01-MAR-2001; 2001WO-US08666.	
XX	PR	09-MAR-2001; 2001US-0802706.	
XX	PR	14-MAR-2001; 2001US-0808689.	
XX	PR	22-MAR-2001; 2001US-0816744.	
XX	PR	05-APR-2001; 2001US-0828366.	
XX	PR	10-MAY-2001; 2001US-0854208.	
XX	PR	25-MAY-2001; 2001US-0866028.	
XX	PR	25-MAY-2001; 2001WO-US17092.	
XX	PR	30-MAY-2001; 2001US-0870574.	
XX	PR	30-MAY-2001; 2001WO-US17443.	
XX	PR	01-JUN-2001; 2001WO-US17800.	
XX	PR	20-JUN-2001; 2001WO-US19692.	
XX	PR	28-JUN-2001; 2001WO-US00000.	
XX		(GETH ) GENENTECH INC.	
PA		(BAKE/) BAKER K P.	
PA		(FERE/) FERRARA N.	
PA		(GERB/) GERBER H.	
PA		(GERR/) GERRITSEN M E.	
PA		(GODD/) GODDARD A.	
PA		(GODO/) GODOWSKI P J.	
PA		(GURN/) GURNEY A L.	
PA		(HILL/) HILLAN K J.	

PA (MARS/) MARSTERS S A.  
PA (PANJ/) PAN J.  
PA (PAON/) PAONI N F.  
PA (STEP/) STEPHAN J F.  
PA (WATA/) WATANABE C K.  
PA (WILL/) WILLIAMS F M.  
PA (WOOD/) WOOD W I.  
XX Baker KP, Ferrata N, Gerber H, Gerritsen ME, Goddard A;  
PI Godowski PJ, Gurney AL, Hillan KJ, Marsters SA, Pan J, Paoni NF;  
PI Stephan JF, Watanabe CK, Williams FM, Wood WI, Ye W;  
XX WPI: 2002-171999/22.  
DR N-PSDB; ABU95597.  
XX  
PT One hundred and eighty seven nucleic acids encoding PRO polypeptides,  
PT useful in diagnosis and treatment of cardiovascular (e.g. myocardial  
PT infarction), endothelial or angiogenic disorders in a mammal -  
XX  
XX Claim 11; Fig 74; 567pp; English.  
XX  
CC The present invention provides the protein and coding sequences of human  
CC PRO proteins. These are useful for treating or diagnosing a  
CC cardiovascular, endothelial or angiogenic disorder, including cardiac  
CC hypertrophy, trauma, cancer, age-related macular degeneration,  
CC atherosclerosis, hypertension, arterial restenosis, rheumatoid arthritis,  
CC angina, myocardial infarctions, thrombophlebitis, lymphangitis, tumour  
CC angiogenesis (such as breast carcinoma and liver carcinoma) and wound  
CC healing. The present sequence is a PRO protein of the invention.  
XX  
XX Sequence 867 AA;  
XX

Query Match 99.3%; Score 4719.5; DB 23; Length 867;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 866; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1 MGPPSLVLCCLLSATVFSLLGSSAFSLSHRLKGRFORRRIRPNILVLTDQDVELGS 60  
DB 1 MGPPSLVLCCLLSATVFSLLGSSAFSLSHRLKGRFORRRIRPNILVLTDQDVELGS 60  
QY 61 MQVMNKTREIMEQGGHFNFAVTTTMCPCSSSILTKGVYVHNHTYNNENCSPSQWA 120  
DB 61 MQVMNKTREIMEQGGHFNFAVTTTMCPCSSSILTKGVYVHNHTYNNENCSPSQWA 120  
QY 121 QHESRTFAVYLNSTGYRTAFCKYLYNEVNGSVYPPGKWKVGLLNKSPFYNTLCRNGVK 180  
DB 121 QHESRTFAVYLNSTGYRTAFCKYLYNEVNGSVYPPGKWKVGLLNKSPFYNTLCRNGVK 180  
QY 181 EXHGSYSKDYLTDLITNDSVSFFRTSKQNYPHRPVLMVISHAAPHGPEDSAPQYSRLFP 240  
DB 181 EXHGSYSKDYLTDLITNDSVSFFRTSKQNYPHRPVLMVISHAAPHGPEDSAPQYSRLFP 240  
QY 241 NASQHTTPSYNAPNPDGHWIMRYTGPMPKPIHMETNMLQKRLQTLMSVDDSMETIYNN 300  
DB 241 NASQHTTPSYNAPNPDGHWIMRYTGPMPKPIHMETNMLQKRLQTLMSVDDSMETIYNN 300  
QY 301 LVETGELDNTIYVYTAADHGHIHQFGLVKGKSMPEYFDIRVFPYVGNVVEAGCLNPHIV 360  
DB 301 LVETGELDNTIYVYTAADHGHIHQFGLVKGKSMPEYFDIRVFPYVGNVVEAGCLNPHIV 360  
QY 361 LNIDLAFTILDIAGLDIPADMGKSLKLLDTERPVRNPHLKKQVRWRDSDLVERGKLL 420  
DB 361 LNIDLAFTILDIAGLDIPADMGKSLKLLDTERPVRNPHLKKQVRWRDSDLVERGKLL 420  
QY 421 HKRDNDKVAQENFUPKYQVRKDLQCORAEYQTAQEQLGKQKQCVEDATGKLKHKCKGP 480  
DB 421 HKRDNDKVAQENFUPKYQVRKDLQCORAEYQTAQEQLGKQKQCVEDATGKLKHKCKGP 480  
QY 481 MRLGGSRLSNLVPKYGGGSEACTCDSDGYKLSLAGRRKKLFKKKYKASYVRSRSRSV 540  
DB 481 MRLGGSRLSNLVPKYGGGSEACTCDSDGYKLSLAGRRKKLFKKKYKASYVRSRSRSV 540  
QY 541 AIEVDGRVYHVLGDAAPRNLTKRHWPGAPEDQDDKGGDFSGTGLPDYSAANPIKVT 600

DB 541 AIEVDGRVYHVLGDAAPRNLTKRHWPGAPEDQDDKGGDFSGTGLPDYSAANPIKVT 600  
QY 601 HRCYILENDTVQCDDLYKSLQAMKDHKLHIDHEITLQNKIKNLREVRGHLKKRPEEC 660  
DB 601 HRCYILENDTVQCDDLYKSLQAMKDHKLHIDHEITLQNKIKNLREVRGHLKKRPEEC 660  
QY 661 DCHKISYTHQHKRGLKHRGSSSLHFFRKGLOEKQKVVLLREQKRKKLRLKRLQNNDTG 720  
DB 661 DCHKISYTHQHKRGLKHRGSSSLHFFRKGLOEKQKVVLLREQKRKKLRLKRLQNNDTG 720  
QY 721 SMPGLTCTHNDQHWOTAPFWTLGPFCACTISANNNTYWCMTINETNPFCEPATGFL 780  
DB 721 SMPGLTCTHNDQHWOTAPFWTLGPFCACTISANNNTYWCMTINETNPFCEPATGFL 780  
QY 781 YFDLNTDPYQLMNAVNTLDRDLNQLHVLQVLMELRSCKGYKQCNPRTRNMOLGLKDGGSYE 840  
DB 781 YFDLNTDPYQLMNAVNTLDRDLNQLHVLQVLMELRSCKGYKQCNPRTRNMOLGLKDGGSYE 840  
QY 841 QYRQQRKRWPEMKRPPSSKSLGQLWEGWEG 870  
DB 838 QYRQQRKRWPEMKRPPSSKSLGQLWEGWEG 867

RESULT 15  
ABB84853  
ID ABB84853 standard; Protein; 867 AA.  
XX ABB84853;  
XX  
XX 16-MAY-2002 (first entry)  
XX  
DE Human PRO1120 protein sequence SEQ ID NO:74.  
XX  
KW Human; angiogenesis; cardiant; cytostatic; antiangiogenic; hypotensive;  
KW vulnerary; antiarteriosclerotic; PRO agonist; PRO antagonist; trauma;  
KW gene therapy; cardiovascular disorder; endothelial disorder; cancer;  
KW angiogenic disorder; cardiac hypertrophy; atherosclerosis; hypertension;  
KW age-related macular degeneration; arterial restenosis; angina;  
KW rheumatoid arthritis; myocardial infarction; thrombophlebitis;  
KW lymphangitis; tumour angiogenesis; breast carcinoma; liver carcinoma;  
KW wound healing; chromosome mapping; gene mapping.  
XX  
OS Homo sapiens.  
XX  
XX WO200200690-A2.  
XX  
XX 03-JAN-2002.  
XX  
XX 20-JUN-2001; 2001WO-US19692.  
XX  
XX 23-JUN-2000; 2000US-213637P.  
XX 20-JUL-2000; 2000US-219556P.  
XX 25-JUL-2000; 2000US-220824P.  
XX 25-JUL-2000; 2000US-220664P.  
XX 28-JUL-2000; 2000WO-US20710.  
XX 02-AUG-2000; 2000US-222695P.  
XX 17-AUG-2000; 2000US-0643657.  
XX 23-AUG-2000; 2000WO-US23522.  
XX 24-AUG-2000; 2000WO-US23328.  
XX 07-SEP-2000; 2000US-230978P.  
XX 18-SEP-2000; 2000US-0664610.  
XX 18-SEP-2000; 2000US-0665350.  
XX 24-OCT-2000; 2000US-242922P.  
XX 08-NOV-2000; 2000US-0709238.  
XX 08-NOV-2000; 2000WO-US30952.  
XX 10-NOV-2000; 2000WO-US30873.  
XX 01-DEC-2000; 2000WO-US32678.  
XX 20-DEC-2000; 2000US-0747259.  
XX 20-DEC-2000; 2000WO-US34956.  
XX 23-JAN-2001; 2001US-0767609.  
XX 28-FEB-2001; 2001US-0796498.  
XX 28-FEB-2001; 2001WO-US06520.



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2003, 13:02:45 ; Search time 54 Seconds  
(without alignments)  
4157.514 Million cell updates/sec

Title: US-09-970-287-2

Perfect score: 4751  
Sequence: 1 MGPPSLVLCLSATVFSLLG.....PMKRPSSKSLQLWEGWG 870

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL 23:

- 1: sp\_archaea:
- 2: sp\_bacteria:
- 3: sp\_fungi:
- 4: sp\_human:
- 5: sp\_invertebrate:
- 6: sp\_mammal:
- 7: sp\_mhc:
- 8: sp\_organelle:
- 9: sp\_phage:
- 10: sp\_plant:
- 11: sp\_rodent:
- 12: sp\_virus:
- 13: sp\_vertebrate:
- 14: sp\_unclassified:
- 15: sp\_virus:
- 16: sp\_bacteriap:
- 17: sp\_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	4751	100.0	885	Q9ULH3	Q9ulh3 homo sapien
2	4746	99.9	870	Q8IWU5	Q8iwu5 homo sapien
3	4469.5	94.1	875	Q8CFG0	Q8cfq0 mus musculu
4	4466.5	94.0	875	Q8BZ51	Q8bz51 mus musculu
5	4466.5	94.0	908	Q8BM68	Q8bm68 mus musculu
6	4462.5	93.9	875	Q8BU24	Q8bu24 mus musculu
7	4458.5	93.8	875	Q8C169	Q8cl69 mus musculu
8	3175.5	66.8	871	Q8IWU6	Q8iwu6 homo sapien
9	3142	66.1	870	Q8V160	Q8v160 rattus norv
10	3131.5	65.9	870	Q8K007	Q8k007 mus musculu
11	3128.5	65.8	870	Q8CLD3	Q8cl3 mus musculu
12	3117.5	65.6	867	Q90XB6	Q90xb6 coturnix co
13	3094.5	65.1	862	Q8BLJ0	Q8blj0 mus musculu
14	3059.5	64.4	604	Q8BX28	Q8bx28 mus musculu
15	3059	64.4	818	Q9UPS5	Q9ups5 homo sapien
16	2827	59.5	525	Q96SG2	Q96sg2 homo sapien

17	1914	40.3	344	4	Q9H1G9	Q9hl9 homo sapien
18	1714.5	36.1	1114	5	Q9VEX0	Q9vex0 drosophila
19	1681	35.4	540	4	Q8NCA2	Q8nca2 homo sapien
20	1589	33.4	709	5	Q2L376	Q2l376 caenorhabdi
21	1471.5	31.0	1006	5	Q8MRG1	Q8mrg1 drosophila
22	1420	29.9	983	5	Q9N1H6	Q9n1h6 drosophila
23	912	19.2	547	13	Q8AVH8	Q8avh8 xenopus lae
24	869.5	18.3	544	11	Q8BJJ7	Q8bjj7 mus musculu
25	868.5	18.3	544	11	Q8BK91	Q8bk91 mus musculu
26	867.5	18.3	544	11	Q8BFR4	Q8bfr4 mus musculu
27	855	18.0	149	4	Q9JTR3	Q9jtr3 homo sapien
28	831	17.5	149	11	Q3DS82	Q3ds82 mus musculu
29	812	17.1	427	6	Q8MJB8	Q8mjb8 equus cabal
30	702	14.8	124	4	Q9H1H0	Q9h1h0 homo sapien
31	679	14.3	492	5	Q9I7F8	Q9i7f8 drosophila
32	674	14.2	492	5	Q9SR73	Q9sr73 drosophila
33	453.5	9.5	639	3	Q43113	Q43113 neurospora
34	430.5	9.1	649	10	Q9ATG5	Q9atg5 chlamydomon
35	364	7.7	517	2	Q9L5W0	Q9l5w0 prevotella
36	323	6.8	511	3	Q9P8C2	Q9p8c2 colicortic
37	301.5	6.3	497	16	Q8XC09	Q8xc09 escherichia
38	298.5	6.3	497	16	Q8P8W9	Q8pbw9 escherichia
39	291.5	6.1	169	5	Q8MVL7	Q8mvl7 boltenia vi
40	290	6.1	535	16	Q8ZHP8	Q8zhp8 versinia pe
41	279.5	5.9	491	11	Q8BL50	Q8bl50 mus musculu
42	279.5	5.9	556	11	Q9D2L1	Q9d2l1 mus musculu
43	279.5	5.9	556	11	Q8BUA1	Q8bua1 mus musculu
44	279.5	5.9	563	11	Q9CYZ0	Q9cyz0 mus musculu
45	271	5.7	562	5	Q8SZ72	Q8sz72 drosophila

## ALIGNMENTS

## RESULT 1

Q9ULH3 ID Q9ULH3 PRELIMINARY; PRT; 885 AA.  
AC Q9ULH3;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Hypothetical protein KIAA1247 (Fragment).  
GN KIAA1247.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=20039619; PubMed=10574462;  
RA Nagase T., Ishikawa K., Kikuno R., Hirosewa M., Nomura N., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. XV.  
RT The complete sequences of 100 new cDNA clones from brain which code  
RT for large proteins in vitro."  
RL DNA Res. 6:337-345 (1999).  
DR EMBL; AB033073; BAA86561.2; .  
DR InterPro; IPR002129; Pyridoxal dec.  
DR Pfam; PF00884; Sulfatase\_1.  
DR PROSITE; PS00392; DDC\_GAD\_HDC\_YDC; 1.  
DR PROSITE; PS00523; SULFATASE\_1; 1.  
FT NON TER 1  
SQ SEQUENCE 885 AA; 102234 MW; CCC1D4A8297AEE16 CRC64;

Query Match 100.0%; Score 4751; DB 4; Length 885;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGPPSLVLCLSATVFSLLGSSAFSLSHRLKGRFQDRNRIRPNILVLTDDQDVELGS 60

Db 16 MGPPSLVLCLSATVFSLLGSSAFSLSHRLKGRFQDRNRIRPNILVLTDDQDVELGS 75



```

QY 61 MQVWNTTRIMEGGTHFINAFVTTIMCCPSRSSILTKGYVHNHTYTNNECSSPSWQA 120
DB 76 MQVWNTTRIMEGGTHFINAFVTTIMCCPSRSSILTKGYVHNHTYTNNECSSPSWQA 135
QY 121 QHESRTFAVYLNSTGVTAFPGKYLYNEYSYVPPGKSWGLLNKSRFYNTYTLCRNGVK 180
DB 136 QHESRTFAVYLNSTGVTAFPGKYLYNEYSYVPPGKSWGLLNKSRFYNTYTLCRNGVK 195
QY 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAHPGSDSAPQYSRLFP 240
DB 196 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAHPGSDSAPQYSRLFP 255
QY 241 NASQHTPTSYNAPNDKHWIMRYTGMKPIHMEFTNMLQRLQTLMSVDDSMETIYNN 300
DB 256 NASQHTPTSYNAPNDKHWIMRYTGMKPIHMEFTNMLQRLQTLMSVDDSMETIYNN 315
QY 301 LVETGELDNTIYVYTADHGHIQOGLVKGSMPYBFDIRVPYVPGPNVAGCLNPHIV 360
DB 316 LVETGELDNTIYVYTADHGHIQOGLVKGSMPYBFDIRVPYVPGPNVAGCLNPHIV 375
QY 361 LNTDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKQVWRDSSFVERGKLL 420
DB 376 LNTDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKQVWRDSSFVERGKLL 435
QY 421 HKRDNDKVDQAEENFLPKYQVRKDLQRAEYOTACEOLGQKQWCVEDATGKLLHKCKGP 480
DB 436 HKRDNDKVDQAEENFLPKYQVRKDLQRAEYOTACEOLGQKQWCVEDATGKLLHKCKGP 495
QY 481 MRLGGSRALSNLVPKYVGGSEACTCDSGDYKLSLAGRRKKLFKKYKASYVRSRSIRSV 540
DB 496 MRLGGSRALSNLVPKYVGGSEACTCDSGDYKLSLAGRRKKLFKKYKASYVRSRSIRSV 555
QY 541 AIEVDGRVYHVLGDAAQAPNLTKRHWPGAPEDQDDKGGDFSGTGLPDYSAANPIKVT 600
DB 556 AIEVDGRVYHVLGDAAQAPNLTKRHWPGAPEDQDDKGGDFSGTGLPDYSAANPIKVT 615
QY 601 HRCYILENDTVQCDLDLYKSLQAWKDKHLHIDHEIETLQNKIKNLRVGRHLKKKPEEC 660
DB 616 HRCYILENDTVQCDLDLYKSLQAWKDKHLHIDHEIETLQNKIKNLRVGRHLKKKPEEC 675
QY 661 DCHKISYHTQHKGRHLKXRGSSLHPPFRKGLQEKDKVWLLREQKFKKLRKLLKRLQNNDT 720
DB 676 DCHKISYHTQHKGRHLKXRGSSLHPPFRKGLQEKDKVWLLREQKFKKLRKLLKRLQNNDT 735
QY 721 SMPGLTCFTHDNQHWOTAPFTWTLGPFCACTSANNTYWCMTINETHNLFCEPATGFLE 780
DB 736 SMPGLTCFTHDNQHWOTAPFTWTLGPFCACTSANNTYWCMTINETHNLFCEPATGFLE 795
QY 781 YFDLNTDPYOLMNAVNTLDRDVLNQLHVQWELRSCGYKQCNPRTRNNMDLGLKDGGSYE 840
DB 796 YFDLNTDPYOLMNAVNTLDRDVLNQLHVQWELRSCGYKQCNPRTRNNMDLGLKDGGSYE 855
QY 841 QYRQFORRWKPEMKRPSKSLGOLWEGWEG 870
DB 856 QYRQFORRWKPEMKRPSKSLGOLWEGWEG 885

```

RESULT 2

Q8IWUS

ID Q8IWUS PRELIMINARY; PRT: 870 AA.

AC Q8IWUS;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Extracellular sulfatase SULF-2.

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Butheria; Primates; Catarrhini; Hominiidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Lung;

```

RX PubMed=12368295;
RA Morimoto-Tomita M., Uchimura K., Werb Z., Hemmerich S., Rosen S.D.;
RT "Cloning and Characterization of Two Extracellular Heparin-degrading
RT Endosulfatases in Mice and Humans.";
EL J. Biol. Chem. 277:49175-49185 (2002).
DR EMBL; AY101176; AA076861.1; -. 74B1069CE2774D73 CRC64;
SQ SEQUENCE 870 AA; 100434 MW; 74B1069CE2774D73 CRC64;

```

Query Match 99.9%; Score 4746; DB 4; Length 870;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MGPPSLVLCLLSATVFSLLGGSSAFISHRLKGRFORDRRNIRPNIIILVTDQDVELGS 60
DB 1 MGPPSLVLCLLSATVFSLLGGSSAFISHRLKGRFORDRRNIRPNIIILVTDQDVELGS 60
QY 61 MQVWNTTRIMEGGTHFINAFVTTIMCCPSRSSILTKGYVHNHTYTNNECSSPSWQA 120
DB 61 MQVWNTTRIMEGGTHFINAFVTTIMCCPSRSSILTKGYVHNHTYTNNECSSPSWQA 120
QY 121 QHESRTFAVYLNSTGVTAFPGKYLYNEYSYVPPGKSWGLLNKSRFYNTYTLCRNGVK 180
DB 121 QHESRTFAVYLNSTGVTAFPGKYLYNEYSYVPPGKSWGLLNKSRFYNTYTLCRNGVK 180
QY 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAHPGSDSAPQYSRLFP 240
DB 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAHPGSDSAPQYSRLFP 240
QY 241 NASQHTPTSYNAPNDKHWIMRYTGMKPIHMEFTNMLQRLQTLMSVDDSMETIYNN 300
DB 241 NASQHTPTSYNAPNDKHWIMRYTGMKPIHMEFTNMLQRLQTLMSVDDSMETIYNN 300
QY 301 LVETGELDNTIYVYTADHGHIQOGLVKGSMPYBFDIRVPYVPGPNVAGCLNPHIV 360
DB 301 LVETGELDNTIYVYTADHGHIQOGLVKGSMPYBFDIRVPYVPGPNVAGCLNPHIV 360
QY 361 LNTDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKQVWRDSSFVERGKLL 420
DB 361 LNTDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKQVWRDSSFVERGKLL 420
QY 421 HKRDNDKVDQAEENFLPKYQVRKDLQRAEYOTACEOLGQKQWCVEDATGKLLHKCKGP 480
DB 421 HKRDNDKVDQAEENFLPKYQVRKDLQRAEYOTACEOLGQKQWCVEDATGKLLHKCKGP 480
QY 481 MRLGGSRALSNLVPKYVGGSEACTCDSGDYKLSLAGRRKKLFKKYKASYVRSRSIRSV 540
DB 481 MRLGGSRALSNLVPKYVGGSEACTCDSGDYKLSLAGRRKKLFKKYKASYVRSRSIRSV 540
QY 541 AIEVDGRVYHVLGDAAQAPNLTKRHWPGAPEDQDDKGGDFSGTGLPDYSAANPIKVT 600
DB 541 AIEVDGRVYHVLGDAAQAPNLTKRHWPGAPEDQDDKGGDFSGTGLPDYSAANPIKVT 600
QY 601 HRCYILENDTVQCDLDLYKSLQAWKDKHLHIDHEIETLQNKIKNLRVGRHLKKKPEEC 660
DB 601 HRCYILENDTVQCDLDLYKSLQAWKDKHLHIDHEIETLQNKIKNLRVGRHLKKKPEEC 660
QY 661 DCHKISYHTQHKGRHLKXRGSSLHPPFRKGLQEKDKVWLLREQKFKKLRKLLKRLQNNDT 720
DB 661 DCHKISYHTQHKGRHLKXRGSSLHPPFRKGLQEKDKVWLLREQKFKKLRKLLKRLQNNDT 720
QY 721 SMPGLTCFTHDNQHWOTAPFTWTLGPFCACTSANNTYWCMTINETHNLFCEPATGFLE 780
DB 721 SMPGLTCFTHDNQHWOTAPFTWTLGPFCACTSANNTYWCMTINETHNLFCEPATGFLE 780
QY 781 YFDLNTDPYOLMNAVNTLDRDVLNQLHVQWELRSCGYKQCNPRTRNNMDLGLKDGGSYE 840
DB 781 YFDLNTDPYOLMNAVNTLDRDVLNQLHVQWELRSCGYKQCNPRTRNNMDLGLKDGGSYE 840
QY 841 QYRQFORRWKPEMKRPSKSLGOLWEGWEG 870
DB 841 QYRQFORRWKPEMKRPSKSLGOLWEGWEG 870

```

```

RESULT 3
Q8CFG0 PRELIMINARY; PRT; 875 AA.
ID Q8CFG0;
AC Q8CFG0;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE Extracellular sulfatase SULF-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Gross mammary tumor;
RX PubMed=12368295;
RA Morimoto-Tomita M., Uchimura K., Werb Z., Hemmerich S., Rosen S.D.;
RT "Cloning and Characterization of Two Extracellular Heparin-degrading
RT Endosulfatases in Mice and Humans";
RL J. Biol. Chem. 277:49175-49185 (2002).
DR EMBL: AY101117; AAM76862.1; -.
SQ SEQUENCE 875 AA; 100469 MW; 148406BCCC7D76B9 CRC64;

Query Match 94.1%; Score 4469.5; DB 11; Length 875;
Best Local Similarity 93.1%; Pred. No. 0;
Matches 821; Conservative 18; Mismatches 30; Indels 7; Gaps 3;

QY 1 MGPPSLVCLLSATVFSLLGSSAFSLSHRLKGRFQDRNRIRNPNIILVLTDDQDVELGS 60
Db 1 MAPPGPLWLLSTALLSLLAGSSAFSLSHRLKGRFQDRNRIRNPNIILVLTDDQDVELGS 60
QY 61 MQVMKTRIMEGGTHFINAFVTPPMCCPSRSSILTGKYVHNHTYTNNECSPSQWA 120
Db 61 MQVMKTRIMEGGTHFINAFVTPPMCCPSRSSILTGKYVHNHTYTNNECSPSQWA 120
QY 121 QHESRTFAVILNSTGYRTAFFGKYLNEYNGSVYPPGKWEVGLLKNRFFNYTLCRNGVK 180
Db 121 QHESRTFAVILNSTGYRTAFFGKYLNEYNGSVYPPGKWEVGLLKNRFFNYTLCRNGVK 180
QY 181 EXHGSYSKDYLTDLITNDVSFFRTSKMYPHRLPVLVISHAAPHGEDSAPQYSRLFP 240
Db 181 EXHGSYSKDYLTDLITNDVSFFRTSKMYPHRLPVLVISHAAPHGEDSAPQYSRLFP 240
QY 241 NASQHTPSYVAPNDKHWIMRYTGPMPKIHMFTNMLQKRLQTLMSVDDSMETIYNM 300
Db 241 NASQHTPSYVAPNDKHWIMRYTGPMPKIHMFTNMLQKRLQTLMSVDDSMETIYNM 300
QY 301 LVETGELDNITYIVYTADHGTHIGQFGLVKGSMPEYFDIRVPFYVGPVNEAGSLNPHIV 360
Db 301 LVETGELDNITYIVYTADHGTHIGQFGLVKGSMPEYFDIRVPFYVGPVNEAGSLNPHIV 360
QY 361 LNIDLAPTILDTAGLDIPADMDGKSLILKLLDTERPVRNPHLKKLRVWRDSFLVERGKLL 420
Db 361 LNIDLAPTILDTAGLDIPADMDGKSLILKLLDTERPVRNPHLKKLRVWRDSFLVERGKLL 420
QY 421 HKRDNDKVAQENFLPKYQVVDLQCORAEYQTACEQLGQKQVCVEDATGKLKHKCKGP 480
Db 421 HKRDNDKVAQENFLPKYQVVDLQCORAEYQTACEQLGQKQVCVEDATGKLKHKCKGP 480
QY 481 MRL- --GGSRALSNLPKYGGGSEACTCDS- --GDYKLSLAGRRKLPKKYKASVYRS 534
Db 481 MRL- --GGSRALSNLPKYGGGSEACTCDS- --GDYKLSLAGRRKLPKKYKASVYRS 534
QY 535 RSIRSAIEVDGRVYHVLGDAAPRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAA 594
Db 535 RSIRSAIEVDGRVYHVLGDAAPRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAA 594
QY 540 RSIRSAIEVDGEIYHVGLDTPQPRNLSKPHWPGAPEDQDDKGGDFSGTGGLPDYSAP 599
Db 540 RSIRSAIEVDGEIYHVGLDTPQPRNLSKPHWPGAPEDQDDKGGDFSGTGGLPDYSAP 599
QY 595 NP1KVTNRCVILENDTVQCDLDLYKSLQAKDKHLHIHEIETLQNK1KNLREVRGHLK 654
Db 595 NP1KVTNRCVILENDTVQCDLDLYKSLQAKDKHLHIHEIETLQNK1KNLREVRGHLK 654
QY 600 NP1KVTNRCVILENDTVQCDLDLYKSLQAKDKHLHIHEIETLQNK1KNLREVRGHLK 659
Db 600 NP1KVTNRCVILENDTVQCDLDLYKSLQAKDKHLHIHEIETLQNK1KNLREVRGHLK 659
QY 655 KRPEECDCHKISYTHQKRLKRGSSLPFRKGLQEKDKVLLRQEKKKLKLKRL 714
Db 655 KRPEECDCHKISYTHQKRLKRGSSLPFRKGLQEKDKVLLRQEKKKLKLKRL 714

RESULT 4
Q8BZ51 PRELIMINARY; PRT; 875 AA.
ID Q8BZ51;
AC Q8BZ51;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE Weakly similar to N-acetylglucosamine-6-sulfatase.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Bone;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL: AK036685; BAC22934.1; -.
SQ SEQUENCE 875 AA; 100497 MW; A04AB401029FCCDC CRC64;

Query Match 94.0%; Score 4466.5; DB 11; Length 875;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 820; Conservative 19; Mismatches 30; Indels 7; Gaps 3;

QY 1 MGPPSLVCLLSATVFSLLGSSAFSLSHRLKGRFQDRNRIRNPNIILVLTDDQDVELGS 60
Db 1 MAPPGPLWLLSTALLSLLAGSSAFSLSHRLKGRFQDRNRIRNPNIILVLTDDQDVELGS 60
QY 61 MQVMKTRIMEGGTHFINAFVTPPMCCPSRSSILTGKYVHNHTYTNNECSPSQWA 120
Db 61 MQVMKTRIMEGGTHFINAFVTPPMCCPSRSSILTGKYVHNHTYTNNECSPSQWA 120
QY 121 QHESRTFAVILNSTGYRTAFFGKYLNEYNGSVYPPGKWEVGLLKNRFFNYTLCRNGVK 180
Db 121 QHESRTFAVILNSTGYRTAFFGKYLNEYNGSVYPPGKWEVGLLKNRFFNYTLCRNGVK 180
QY 181 EXHGSYSKDYLTDLITNDVSFFRTSKMYPHRLPVLVISHAAPHGEDSAPQYSRLFP 240
Db 181 EXHGSYSKDYLTDLITNDVSFFRTSKMYPHRLPVLVISHAAPHGEDSAPQYSRLFP 240
QY 241 NASQHTPSYVAPNDKHWIMRYTGPMPKIHMFTNMLQKRLQTLMSVDDSMETIYNM 300
Db 241 NASQHTPSYVAPNDKHWIMRYTGPMPKIHMFTNMLQKRLQTLMSVDDSMETIYNM 300
QY 301 LVETGELDNITYIVYTADHGTHIGQFGLVKGSMPEYFDIRVPFYVGPVNEAGSLNPHIV 360
Db 301 LVETGELDNITYIVYTADHGTHIGQFGLVKGSMPEYFDIRVPFYVGPVNEAGSLNPHIV 360
QY 361 LNIDLAPTILDTAGLDIPADMDGKSLILKLLDTERPVRNPHLKKLRVWRDSFLVERGKLL 420
Db 361 LNIDLAPTILDTAGLDIPADMDGKSLILKLLDTERPVRNPHLKKLRVWRDSFLVERGKLL 420
QY 421 HKRDNDKVAQENFLPKYQVVDLQCORAEYQTACEQLGQKQVCVEDATGKLKHKCKGP 480
Db 421 HKRDNDKVAQENFLPKYQVVDLQCORAEYQTACEQLGQKQVCVEDATGKLKHKCKGP 480

```

421	HKRGDKVNAQENFLPKYQKVXKOLCORAEYQVTACEQLGQKMQCQVEDASGTLKLHKKGP	480
481	MRLL--CGSRALSNLVPKYQGSSEACTCDS--GDYKLSLAGRRKKLKKKYKASYVRS	534
481	MRFGGGGSRALSNLVPKYDQGSSEACSDSGGGDYKLAGRR-KLKKKYKTSYARN	539
535	RSIRSAIEVGRVYHVGIGDAAPRNLTKRHPGAPEDODDKGDDPFGTGLPDYGA	594
540	RSIRSAIEVDEGEIYHVGIDTVPOPRNLSPHPMPGAPEDODDKGDDPFGTGLPDYSA	599
595	NP1KVTHRCYILENDTVOCDDLDLKYSLQAKWDHKLHIDHIEITLQNKIKNLREVRGHLK	654
600	NP1KVTHRCYILENDTVOCDDLDLKYSLQAKWDHKLHIDHIEITLQNKIKNLREVRGHLK	659
655	KRPECDCHKISYHTQHKGRLLKRGSSLHPPRGLQEQKVKVLLRERQKKLRLKLL	714
660	KRPECDCHRISYHSQHKGRLLKHKGSSLHPPRGLQEQKVKVLLRERQKKLRLKLL	719
715	QNNDTCSMPGLTCFTHDHHQHOTAPPTLGPFCACISANNNTYCHRTINETHNFICEP	779
720	QNNDTCSMPGLTCFTHDHHQHOTAPLTLGPFCACISANNNTYCHRTINETHNFICEP	779
775	ATGFEYFDLNTDPYQLMNAVNTLDRDVJNLQHVLMELRSCYKQCQCPNRTNMDLGLK	834
780	ATGFEYFDLSTDPYQLMNAVNTLDRDVJNLQHVLMELRSCYKQCQCPNRTNMDLGLR	839
835	DGGSYEYQRFQRRKWPENKRPSSKSLQGLWEGWG	870
840	DGGSYEYQRFQRRKWPENKRPSSKSLQGLWEGWG	875
RESULT 5		
Q8BM68		
ID	Q8BM68	PRELIMINARY; PRT; 908 AA.
AC	Q8BM68;	
DT	01-MAR-2003 (TrEMBLrel. 23, Created)	
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)	
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)	
DE	Weakly similar to N-acetylglucosamine-6-sulfatase.	
OS	Mus musculus (Mouse)	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
OC	NCBI_TaxId=10090;	
RN	[1]	
RF	SEQUENCE FROM N.A.	
RC	STRAIN=C57BL/6J; TISSUE=Embryo;	
RX	MEDLINE=22354683; PubMed=12466851;	
RA	The PANTOM Consortium,	
RA	the RIKEN Genome Exploration Research Group Phase I & II Team;	
RT	"Analysis of the mouse transcriptome based on functional annotation of	
RT	60,770 full-length cDNAs."	
RL	Nature 420:563-573 (2002).	
DR	EMBL; AK034712; BAC28804.1; ..	
SQ	SEQUENCE 908 AA; 104037 MW; 1B460655AEC3734D CRC64;	
Query Match 94.0%; Score 4466.5; DB 11; Length 908;		
Best Local Similarity 93.6%; Pred. No. 0;		
Matches 820; Conservative 19; Mismatches 30; Indels 7; Gaps 3		
QY	1 MGPPSLVCLLSATVFLSLGGSSAFLSHRHRLKGRFQDRNRIRPNILVLTDQDVELGS	60
DB	34 MAPPGLPLWLLSTALLSLLAGSSAFLSHRLKGRFQDRNRIRPNILVLTDQDVELGS	93
QY	61 MQVMNTRIMEGGTHFINAFVTPMCCPSRSSILTKYVHHNTYTNENCSSPSWOA	120
DB	94 MQVMNTRIMEGGGAHFINAFVTPMCCPSRSSILTKYVHHNTYTNENCSSPSWOA	153
QY	121 QHSRTFAVILNSTGYRTAFGKYLNEYNGSVPPGKWEVGLLQNSRFNYTLCRNGVK	180
DB	154 QHSRTFAVILNSTGYRTAFGKYLNEYNGSVPPGKWEVGLLQNSRFNYTLCRNGVK	213
QY	181 EKHGSDYSKDYLTDLITNDSVFFRTSKQYPHRPVLTMLVSHAAPGPDSPAYQSRLEP	240

Db	214	EXHGSYDSTYLTDLITNDVSFFSTSKVYPHRPVLWVISHAAPHGPDSPAPQYSRLFP	277
Qy	241	NASOHITPSYVAPNPKDGHIMRTYGTMPKPIHMEFTNMLQRKQLTMSYDDSMETIYAM	300
Db	274	NASOHITPSYVAPNPKDGHIMRTYGTMPKPIHMEFTNMLQRKQLTMSYDDSMETIYAM	333
Qy	301	LVTGELONTYIVYTADGHYHIGFGLVKGKSNPYEPDIRVPYVYRGPNVZAGLNPHTV	360
Db	334	LVTGELONTYIVYTADGHYHIGFGLVKGKSNPYEPDIRVPYVYRGPNVZAGLNPHTV	393
Qy	361	LNIDLAPITLIDAGLIDTADMDGKSLKLLADTERPVNRFHLKKKRVWRDSSFIVERGKLL	420
Db	394	LNIDLAPITLIDAGLIDTADMDGKSLKLLADTERPVNRFHLKKKRVWRDSSFIVERGKLL	453
Qy	421	HKRDNDKYDAQBEENFLPKYQRYKDLQORAEYQTACEOLGQKQWCVEDATGKLLHKCKGP	480
Db	454	HKRGDKYNAQBEENFLPKYQRYKDLQORAEYQTACEOLGQKQWCVEDATGKLLHKCKGP	513
Qy	481	MRL---GGSRALSNLPKYGSGSEACTDS---GDYKLSLAGRRKKLPKKYKASVYRS	534
Db	514	MRPGGGGSRALSNLPKYGSGSEACTDS---GDYKLSLAGRRKKLPKKYKASVYRS	572
Qy	535	RSIRSAIEVDEGRVYHVLGDAAPRNLTKRHHMFGAPEDQDDXGDDFSGTGLPDYSAA	594
Db	573	RSIRSAIEVDEGEIYHVLGDTVPQPNLSKPHMFGAPEDQDDXGDDFSGTGLPDYSAP	632
Qy	595	NPDKYTHRCYILENDTVQCDLILYKSLQAKWDHKLHIDHIEITLONKIKNLRVGRHLKK	654
Db	633	NPDKYTHRCYILENDTVQCDLILYKSLQAKWDHKLHIDHIEITLONKIKNLRVGRHLKK	692
Qy	655	KRPEBCDCHKISYHTQHKGRKLKRGSSILHPFRKGLQEKQKVVLLRQKRRKKLRLKEL	714
Db	693	KRPEBCDCHRIYSHSQHGRKLKRGSSILHPFRKGLQEKQKVVLLRQKRRKKLRLKEL	752
Qy	715	QNNDTCSMPGLTCFTHDNQHWQTAFTWTLAPFCACTSANNNTYTWMTNETNPLFCBF	774
Db	753	QNNDTCSMPGLTCFTHDNHWWQTAFTWTLAPFCACTSANNNTYTWMTNETNPLFCBF	812
Qy	775	ATGFLEYFDLNTDYPQLANNAVNTLDROVLNQLHVMELRSCYKQCNPRTRNMDGLK	834
Db	813	ATGFLEYFDLSTDPQLANNAVNTLDROVLNQLHVMELRSCYKQCNPRTRNMDGLR	872
Qy	835	DGGSYEYQRFQRRKWPMPKPSKSLGQLWEGWEG	870
Db	873	DGGSYEYQRFQRRKWPMPKPSKSLGQLWEGWEG	908
RESULT 6			
Q8BUZ4	ID	Q8BUZ4	PRELIMINARY; PRT; 875 AA.
AC	DT	01-MAR-2003 (TREMBLrel. 23, Created)	
DT	DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)	
DT	DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)	
DE	OS	Weakly similar to N-acetylglucosamine-6-sulfatase.	
DE	OS	Mus musculus (Mouse)	
OC	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
NCBI_TaxID	NCBI_TaxID	10090;	
RP	RP	SEQUENCE FROM N.A.	
RC	RC	STRAIN=C57BL/6J; TISSUE=Head;	
EX	EX	MEDLINE=22354683; PubMed=12466851;	
RA	RA	The FANTOM Consortium,	
RA	RA	the RIKEN Genome Exploration Research Group Phase I & II Team;	
RT	RT	"Analysis of the mouse transcriptome based on functional annotation of	
RT	RT	60,770 full-length cDNAs."	
RL	RL	Nature 420:563-573 (2002).	
DR	DR	ENBL; AX081643; BAC38279.1; -;	
SQ	SEQUENCE	875 AA; 100531 MW; 4D05AC0923B026F4 CRC64;	
Query Match			
Best Local Similarity			
93.9%; Score 4462.5; DB 11; Length 875;			
93.5%; Pred. NO. 0;			

Query Match 93.9%; Score 4462.5; DB 11; Length 875;  
Best Local Similarity 93.5%; Pred. No. 0;

Matches		819;	Conservative	19;	Mismatches	31;	Indels	7;	Gaps	3;
QY	1	MGPPSLVCLLSATVFSLLGSSAFLSHRLKGRFORDRNRIRPNII	ILVLTDDQDVELGS	60						
DB	1	MAPPGLPLWLLSTALLSLLAGSSAFLSHRLKGRFORDRNRIRPNII	ILVLTDDQDVELGS	60						
QY	61	MQVMNTRIMEQGGTHFINAFVTTPMCPCSSSILTGKYVHNHTYNNENCS	SPSQA	120						
DB	61	MQVMNTRIMEQGAHFINAFVTTPMCPCSSSILTGKYVHNHTYNNENCS	SPSQA	120						
QY	121	QHSRTFAVYLNSTGYRTAFGKYLYNEVNGSVYPPGKWEVGLLKNR	FYNTLCRNGVK	180						
DB	121	QHSRTFAVYLNSTGYRTAFGKYLYNEVNGSVYPPGKWEVGLLKNR	FYNTLCRNGVK	180						
QY	181	EXHGSYSXDYLTDLTINDSVSFFRTSKMYPHRPVLMVISHAAPHG	PDAPQYSRLFP	240						
DB	181	EXHGSYSXDYLTDLTINDSVSFFRTSKMYPHRPVLMVISHAAPHG	PDAPQYSRLFP	240						
QY	241	NASQHTPSVYAPNPKHIMRYTGPMKPIHMEFTNMLQRLQTLMS	VDDSMETIYNM	300						
DB	241	NASQHTPSVYAPNPKHIMRYTGPMKPIHMEFTNMLQRLQTLMS	VDDSMETIYNM	300						
QY	301	LVETGELDNTIYLTADHGHIHQFGLVKGSMPEYFDIRVPPYV	RGNVAGCLNPHIV	360						
DB	301	LVETGELDNTIYLTADHGHIHQFGLVKGSMPEYFDIRVPPYV	RGNVAGCLNPHIV	360						
QY	361	LNIDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRFHLKXK	VRWDSFLVERGKLL	420						
DB	361	LNIDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRFHLKXK	VRWDSFLVERGKLL	420						
QY	421	HKRDNDKVAQENFLPKYQVVKDLCORAEYQTAQEQLGQKWC	QVEDATGKLHKCKGP	480						
DB	421	HKRDNDKVAQENFLPKYQVVKDLCORAEYQTAQEQLGQKWC	QVEDATGKLHKCKGP	480						
QY	481	MRL---GGSRALSNLVPKYGGSEACTCDS---GDYKLSLAGR	RKXLFKKYKASYR	534						
DB	481	MRL---GGSRALSNLVPKYGGSEACTCDS---GDYKLSLAGR	RKXLFKKYKASYR	534						
QY	535	RSRVAIEVDGRIYHVGLDTPVQPRNLSKPHWPGAPEDDDK	GGDPSTGGLPDYSA	594						
DB	535	RSRVAIEVDGRIYHVGLDTPVQPRNLSKPHWPGAPEDDDK	GGDPSTGGLPDYSA	594						
QY	595	NP1KVTHRCYILENDTVQCDLKYSLQAWKHKLHIDHEIETL	QNK1KNLREVRGHLK	654						
DB	595	NP1KVTHRCYILENDTVQCDLKYSLQAWKHKLHIDHEIETL	QNK1KNLREVRGHLK	654						
QY	655	KRPECDCCHKISYTHQKGRKHGSSLHPFRKGLQEKDKVLL	REQRKXKGLKRL	714						
DB	655	KRPECDCCHKISYTHQKGRKHGSSLHPFRKGLQEKDKVLL	REQRKXKGLKRL	714						
QY	715	QNDTCSMPGLTCTFDHNOHWQTAFTWILGPFCACTSANNTY	WCMTINETHNPLFCE	774						
DB	715	QNDTCSMPGLTCTFDHNOHWQTAFTWILGPFCACTSANNTY	WCMTINETHNPLFCE	774						
QY	775	ATGFLYFDLNTDYPQLMAVNTLDRDVLNQLHVMELRSC	KGKQCNPRTNMDLGLK	834						
DB	775	ATGFLYFDLNTDYPQLMAVNTLDRDVLNQLHVMELRSC	KGKQCNPRTNMDLGLK	834						
QY	835	DGGSYEQYQFORRWPMKRPSSKSLGQWEGWEG	870							
DB	835	DGGSYEQYQFORRWPMKRPSSKSLGQWEGWEG	870							
RESULT 7										
Q8C169	PRELIMINARY; PRT; 875 AA.									
AC	Q8C169									
DT	01-MAR-2003 (T8MBLrel. 23, Created)									
DT	01-MAR-2003 (T8MBLrel. 23, Last sequence update)									
DE	01-MAR-2003 (T8MBLrel. 23, Last annotation update)									
OS	Mus musculus (Mouse)									
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									

OC	Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	SEQUENCE FROM M.A.
RP	STRAIN=C57BL/6J; TISSUE=Skin;
RC	MEDLINE=22354683; PubMed=12466851;
RX	The FANTOM Consortium,
RA	"Analysis of the mouse transcriptome based on functional annotation of
RT	60,770 full-length cDNAs."
RL	Nature 420:563-573(2002).
DR	ENBL; AK028874; BAC26165.1; --
SQ	SEQUENCE 875 AA; 100418 MW; 17D61ACB22D24C7C CRC64;
Query March 93.8%; Score 4458.5; DB 11; Length 875;	
Best Local Similarity 93.5%; Pred. No. 0;	
Matches 819; Conservative 19; Mismatches 31; Indels 7; Gaps 3;	
QY	1 MGPPSLVCLLSATVFSLLGSSAFLSHRLKGRFORDRNRIRPNII
DB	1 MAPPGLPLWLLSTALLSLLAGSSAFLSHRLKGRFORDRNRIRPNII
QY	61 MQVMNTRIMEQGGTHFINAFVTTPMCPCSSSILTGKYVHNHTYNNENCS
DB	61 MQVMNTRIMEQGAHFINAFVTTPMCPCSSSILTGKYVHNHTYNNENCS
QY	121 QHSRTFAVYLNSTGYRTAFGKYLYNEVNGSVYPPGKWEVGLLKNR
DB	121 QHSRTFAVYLNSTGYRTAFGKYLYNEVNGSVYPPGKWEVGLLKNR
QY	181 EXHGSYSXDYLTDLTINDSVSFFRTSKMYPHRPVLMVISHAAPHG
DB	181 EXHGSYSXDYLTDLTINDSVSFFRTSKMYPHRPVLMVISHAAPHG
QY	241 NASQHTPSVYAPNPKHIMRYTGPMKPIHMEFTNMLQRLQTLMS
DB	241 NASQHTPSVYAPNPKHIMRYTGPMKPIHMEFTNMLQRLQTLMS
QY	301 LVETGELDNTIYLTADHGHIHQFGLVKGSMPEYFDIRVPPYV
DB	301 LVETGELDNTIYLTADHGHIHQFGLVKGSMPEYFDIRVPPYV
QY	361 LNIDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRFHLKXK
DB	361 LNIDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRFHLKXK
QY	421 HKRDNDKVAQENFLPKYQVVKDLCORAEYQTAQEQLGQKWC
DB	421 HKRDNDKVAQENFLPKYQVVKDLCORAEYQTAQEQLGQKWC
QY	481 MRL---GGSRALSNLVPKYGGSEACTCDS---GDYKLSLAGR
DB	481 MRL---GGSRALSNLVPKYGGSEACTCDS---GDYKLSLAGR
QY	535 RSRVAIEVDGRIYHVGLDTPVQPRNLSKPHWPGAPEDDDK
DB	535 RSRVAIEVDGRIYHVGLDTPVQPRNLSKPHWPGAPEDDDK
QY	595 NP1KVTHRCYILENDTVQCDLKYSLQAWKHKLHIDHEIETL
DB	595 NP1KVTHRCYILENDTVQCDLKYSLQAWKHKLHIDHEIETL
QY	655 KRPECDCCHKISYTHQKGRKHGSSLHPFRKGLQEKDKVLL
DB	655 KRPECDCCHKISYTHQKGRKHGSSLHPFRKGLQEKDKVLL
QY	715 QNDTCSMPGLTCTFDHNOHWQTAFTWILGPFCACTSANNTY
DB	715 QNDTCSMPGLTCTFDHNOHWQTAFTWILGPFCACTSANNTY
QY	775 ATGFLYFDLNTDYPQLMAVNTLDRDVLNQLHVMELRSC
DB	775 ATGFLYFDLNTDYPQLMAVNTLDRDVLNQLHVMELRSC
QY	834 DGGSYEQYQFORRWPMKRPSSKSLGQWEGWEG
DB	834 DGGSYEQYQFORRWPMKRPSSKSLGQWEGWEG

```

QY 835 DGSSEYQYQFORRWKPEMKRPSKSLGQWEGWEG 870
Db 840 DGSSEYQYQFORRWKPEMKRPSKSLGQWEGWEG 875

RESULT 8
Q81WU6 PRELIMINARY; PRT; 871 AA.
AC Q81WU6;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DE Extracellular sulfatase SULF-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RP PubMed=12368295;
RA Morimoto-Tonita M., Uchimura K., Werb Z., Hemmerich S., Rosen S.D.;
RT "Cloning and Characterization of Two Extracellular Heparin-degrading
RT Endosulfatases in Mice and Humans.";
RL J Biol. Chem. 277:49175-49185(2002).
DR EMBL; AY101175; AAW76860.1; -.
SQ SEQUENCE 871 AA; 101026 MW; 9A90ADB280304364 CRC64;

Query Match 66.1%; Score 3175.5; DB 4; Length 871;
Best Local Similarity 64.9%; Pred. No. 1.5e-244;
Matches 578; Conservative 120; Mismatches 142; Indels 51; Gaps 12;

QY 5 SLVCLLSATVFSLLGSSAFSLSHRLKGRFDRRNPRIILVLTDDQDVELGSLQVM 64
Db 7 ALLAVLGT---ELLGSLCSTVSPRFRGRIQOERKNIRPNIIILVLTDDQDVELGSLQVM 63
QY 65 NKTRRIMEOGTHFINAFVTTMCCPSRSSILTGVVHNHTYTNNECSSPSWQAQHE 124
Db 64 NKTRKIMEHGATFINAFVTTMCCPSRSSILTGVVHNHTYTNNECSSPSWQAQHEP 123
QY 125 RTFAVLNLTGYRTAFPGKLYNEINGSYPPGWKEWVGLLKNRFRNYTLCRNGVKEKH 184
Db 124 RTFAVLNLTGYRTAFPGKLYNEINGSYPPGWKEWVGLLKNRFRNYTLCRNGVKEKH 183
QY 185 SDYSKDYLTDLTNDVSPRFTSKMYPHRPVLMVISHAAPHGPDSPAPQSLPNSAQ 244
Db 184 FDYAKDYFTDLITNESINTYFKMSKMPYDFIRVPPFIRGSPVSGIVPQVLNID 243
QY 245 HTPSYNAPNDKHWIMYTGPMKPIHMEFTNMLQRLQTLMSVDDSMETIYNMLVET 304
Db 244 HTPSYNAPNDKHWIMYTGPMKPIHMEFTNMLQRLQTLMSVDDSMETIYNMLVET 303
QY 305 GELDNTYIVYTDHGYHIGQFGLVKGKSMYPEDIRVPPVGRPNVEAGCLNPHIVL 364
Db 304 GELNTYIITADHGYHIGQFGLVKGKSMYPEDIRVPPVGRPNVEAGCLNPHIVL 363
QY 365 LAPTIIDAGLIDIPADMDGSKILKLDTERPNVRPHLKKMRVWRDSFLVERGKLLH 424
Db 364 LAPTIIDAGLIDIPADMDGSKILKLDTERPNVRPHLKKMRVWRDSFLVERGKLLH 423
QY 425 NDYVDAQEENFLKYQVXDLCORAYOTACEOLGQKQCVEDATGKLLKHCKGPMRLG 484
Db 424 ESSKNIQSSNHLKYERVKELCOQARYOTACEOLGQKQCVEDATGKLLKHCKGPMRL 483
QY 485 GSR-ALSNLQPKYVGQSEACTCDSDGYKLSLAGRR-KKLFFK---KYKASYRRSI 537
Db 484 TVRQSTNLTARGFHDKDECCRESGYRASRSQRKQFLNQOTPKYKPRFVHTROT 543
QY 538 RSVIAEVDGRVYHVLGD-----AAQPRNLTKHMFPG---APEDQDDKGG-----DFS 583
Db 544 RSLSEVEGEIYDINLEEBELQVLPRIAKHGDGHGKPRDLOQASSGNGRMLADSS 603
QY 584 GTGGLPDYSAANPKVTHRCYILIENTVQCDDLLYKSLQAWKHLDIHEITETLQNK 643

```

```

Db 604 NAVGPP-----TTVRTHKCFILPNDSTHCERELQSAWAKHKAVIDKEIALQDKIK 658
QY 644 NLREVRGHLKKRPBECCHKISYHTQHGRLKHR--GSSLHPRPKGLQEXD-KVWLLRE 700
Db 659 NLREVRGHLKKRPBECCHKISYHTQHGRLKHR--GSSLHPRPKGLQEXD-KVWLLRE 718
QY 701 -OKRKKKRLKRLKRLQNDTCSMPGLTCTFHDNQWQTAPFWTLGPPCACTSANNTYWC 759
Db 719 NNRERRKKEKRRQKGBECSLPGITCTFHDNQWQTAPFWTLGPPCACTSANNTYWC 778
QY 760 MRTINETHFLCEPATGPLEYFDLNTDYPQLMVAINTLDRDVLNQLHVLQMLRSCGY 819
Db 779 LRTVNEHFLCEPATGPLEYFDLNTDYPQLMVAINTLDRDVLNQLHVLQMLRSCGY 838
QY 820 KQCNPRTRNMDLGLKGGSEYQYQFORRWKPEMKRPSKSLGQWEGWEG 870
Db 839 KQCNPRPKNLDVGNKGGSYDLHR-----GQLWDGWEW 871

RESULT 9
Q8VI60 PRELIMINARY; PRT; 870 AA.
AC Q8VI60;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Sulfatase PP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RP Ohto T., Uchida H., Yamazaki H., Keino-Masu K., Masu M.;
RT "Identification of a novel non-lysosomal sulfatase expressed in the
RT floor plate, choroid plexus, and cartilage.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF230072; NAL71906.1; -.
DR InterPro; IPR000917; Sulfatase.
DR Pfam; PF00884; Sulfatase; 1.
DR PROSITE; PS00523; SULFATASE_1; 1.
SQ SEQUENCE 870 AA; 100859 MW; 037813868C08967F CRC64;

Query Match 66.1%; Score 3142; DB 11; Length 870;
Best Local Similarity 64.3%; Pred. No. 7e-242;
Matches 572; Conservative 124; Mismatches 144; Indels 50; Gaps 11;

QY 5 SLVCLLSATVFSLLGSSAFSLSHRLKGRFDRRNPRIILVLTDDQDVELGSLQVM 64
Db 7 ALLAVLGT---ELLGFCSTWRSQRFGRVQOERKNIRPNIIILVLTDDQDVELGSLQVM 63
QY 65 NKTRIMEOGTHFINAFVTTMCCPSRSSILTGVVHNHTYTNNECSSPSWQAQHE 124
Db 64 NKTRKIMEHGATFINAFVTTMCCPSRSSILTGVVHNHTYTNNECSSPSWQAQHEP 123
QY 125 RTFAVLNLTGYRTAFPGKLYNEINGSYPPGWKEWVGLLKNRFRNYTLCRNGVKEKH 184
Db 124 RTFAVLNLTGYRTAFPGKLYNEINGSYPPGWKEWVGLLKNRFRNYTLCRNGVKEKH 183
QY 185 SDYSKDYLTDLTNDVSPRFTSKMYPHRPVLMVISHAAPHGPDSPAPQSLPNSAQ 244
Db 184 FDYAKDYFTDLITNESINTYFKMSKMPYDFIRVPPFIRGSPVSGIVPQVLNID 243
QY 245 HTPSYNAPNDKHWIMYTGPMKPIHMEFTNMLQRLQTLMSVDDSMETIYNMLVET 304
Db 244 HTPSYNAPNDKHWIMYTGPMKPIHMEFTNMLQRLQTLMSVDDSMETIYNMLVET 303
QY 305 GELDNTYIVYTDHGYHIGQFGLVKGKSMYPEDIRVPPVGRPNVEAGCLNPHIVL 364
Db 304 GELNTYIITADHGYHIGQFGLVKGKSMYPEDIRVPPVGRPNVEAGCLNPHIVL 363
QY 365 LAPTIIDAGLIDIPADMDGSKILKLDTERPNVRPHLKKMRVWRDSFLVERGKLLH 424

```

```

Db 364 LAPTILDIAGLQPSVDGKSLKLDLEKPGNRFRTNKAKIMRDTFLVERGKFLRKE 423
Qy 425 NDKVDAQEENFLPKYORVKDLCORAEYQACELQGWOCVEDATGKLLKHKCKGPMRLG 484
Db 424 ESSKNIQSNHLPKYERVKELCOOARYQACEOPGQNMOCIEDTSGLRIHKCKGPDLL 483
Qy 485 GSRA-LSNLVPKYGGSEACTCDSDYKLSLGRK-KLFKK-----KYKASYVRSRSI 537
Db 484 TVRQANRLYSRGLQDKDKECHRESGYRPSRQKKEQFLRNOGTPKYKPRFVHTRT 543
Qy 538 RSVAEIVDGRVYHVLGD-----AAQPNLTKRHWP-----APEDDDKGGDFSG 584
Db 544 RSLSVFEGEYDINLEEBELQVLPFRSAKHDEGHOGFICHQAAAGDIRNEMLADSN 603
Qy 585 TGLPDYSAANPIKTHRCVILENDTVQCDLDLYKSLQAWKHLDHIEITLQNKIN 644
Db 604 AVGLP-----TVRVTHKCFILPNDTHCERELYQSARAWKHKAYIDKEIEVLQDKIN 658
Qy 645 LREVRGHLKKRPEECCHKISYHTQHG--RLKHGSSLHPFRK-GLOEKD-KVWLLE 700
Db 659 LREVRGHLKKRPEECGQSYNKEKGVKQKSLHLPFKAAAEVDSKLQLFKE 718
Qy 701 QKXKXKRLKLLKRLQNNDCSMPLGTCFTHDNHWTAPFWTLGPFCACTSANNNTYCM 760
Db 719 HRRKKERKEKQKRGEECSLPGLTCFTHDNHWTAPFWNLGSPCACTSSNNNTYCL 778
Qy 761 RTINETHNLFCEFBATGFLYFDLNTDPVQLMNAVNTLDRDLVNLQHLVQLMELRSCGYK 820
Db 779 RTVNETHNFCEFBATGFLYFDLNTDPVQLMNAVNTLDRDLVNLQHLVQLMELRSCGYK 838
Qy 821 QCNPRTRNMDLGLKGGSYEQYQFORRWKPEMKRPSKSLGQWEGWEG 870
Db 839 QCNPRKSLDVGTEGNGYDPR-----GQLWDGWEG 870

```

## RESULT 10

```

Q8K007 PRELIMINARY; PRT; 870 AA.
AC Q8K007;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Similar to sulfatase FP (Extracellular sulfatase SULF-1).
GN AW121680.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX PubMed=12368295;
RA Morimoto-Tomita M., Uchimura K., Werb Z., Hemmerich S., Rosen S.D.;
RT "Cloning and Characterization of Two Extracellular Heparin-degrading
RT Endosulfatases in Mice and Humans.";
RJ J. Biol. Chem. 277:49175-49185 (2002).
DR EMBL; BC034547; AAH34547.1; -
DR EMBL; AY101178; AAH76863.1; -
DR MGD; MGI:2138563; AW121680.
DR InterPro; IPR000917; Sulfatase.
DR Pfam; PF00884; Sulfatase; 1.
DR PROSITE; PS00523; SULFATASE 1; 1.
SQ SEQUENCE 870 AA; 100922 MW; 4A9BE710D7CF4F9D CRC64;

```

Query Match 65.9%; Score 3131.5; DB 11; Length 870;  
 Best Local Similarity 64.2%; Pred. No. 4.8e-241;  
 Matches 571; Conservative 122; Mismatches 150; Indels 47; Gaps 10;

```

Qy 5 SLVCLLSATFVSLLGSSAFLSHRLKGRFORDRNRPNIILVLTDDQDVGLSGMOWM 64
Db 4 SLWALLAVLTGLLGLSCSTVRSQRFGRIGRIQERKNIRPNILVLTDDQDVGLSLQVM 63
Qy 65 NKTRRIMEOGGTHFINAFVTTMCCPSSSIITGKYVHNHTYTNNECSSPSWQAQHS 124
Db 64 NKTRKIMEOGGATFTNAFVTTMCCPSSSMLTGKYVHNHTYTNNECSSPSWQAHEP 123
Qy 125 RTPAVYLTNGTGYRTAFPGKYLYNEYSYVPPGKWEKVVGLLKNSRFRNYVTLCRNGYKEKHG 184
Db 124 RTFAVYLTNGTGYRTAFPGKYLYNEYSYVPPGKWRWGLGKNSRFRNYVTCENGKEKHG 183
Qy 185 SDYSKDYLTDLTNSVSFFRTSKMYPHRPVLMYISHAHPGPDSPAPQYSRLFPNASQ 244
Db 184 FDYAKDYFTDLITNESINYFKMSKMYPHRPIMVYISHAHPGPDSPAPQFSKLYPNASQ 243
Qy 245 HITPSYNAPNPDQHWIMRYTGPMPKPIHMEFTNMLQRKLOTLMSSVDDSMETIYNMLVET 304
Db 244 HITPSYNAPNPDQHWIMQYTGPMPLIHEFTNVLQRKLOTLMSSVDDSVLRYNMLVES 303
Qy 305 GELDNTYIVTTADHGHYHGGFGLVKGKSMPEYFDIRVFFYVRGPNVEAGCLNPHIVLND 364
Db 304 GELDNTYIIVTTADHGHYHGGFGLVKGKSMPEYFDIRVFFYVRGPNVEAGCLNPHIVLND 363
Qy 365 LAPTILDIAGLQPSVDGKSLKLDLEKPGNRFRTNKAKIMRDTFLVERGKFLRKE 424
Db 364 LAPTILDIAGLQPSVDGKSLKLDLEKPGNRFRTNKAKIMRDTFLVERGKFLRKE 423
Qy 425 NDKVDAQEENFLPKYORVKDLCORAEYQACELQGWOCVEDATGKLLKHKCKGPMRLG 484
Db 424 ESSKNIQSNHLPKYERVKELCOOARYQACEOPGQNMOCIEDTSGLRIHKCKGPDLL 483
Qy 485 GSRA-LSNLVPKYGGSEACTCDSDYKLSLGR-KLKFKK---KYKASYVRSRSI 537
Db 484 TVRQANRLYSRGLQDKDKECHRESGYRPSRQKKEQFLRNOGTPKYKPRFVHTRT 543
Qy 538 RSVAEIVDGRVYHVLGD-----AAQPNLTKRHWP-----APEDDDKGGDFSG 584
Db 544 RSLSVFEGEYDINLEEBELQVLPFRSAKHDEGHOGFICHQAAAGDIRNEMLADSN 603
Qy 585 TGLPDYSAANPIKTHRCVILENDTVQCDLDLYKSLQAWKHLDHIEITLQNKIN 644
Db 604 AVGLP-----ATVRVTHKCFILPNDTHCERELYQSARAWKHKAYIDKEIEVLQDKIN 658
Qy 645 LREVRGHLKKRPEECCHKISYHTQHG--RLKHGSSLHPFRK-GLOEKD-KVWLLE 700
Db 659 LREVRGHLKKRPEECGQSYNKEKGVKQKSLHLPFKAAAEVDSKLQLFKE 718
Qy 701 QKXKXKRLKLLKRLQNNDCSMPLGTCFTHDNHWTAPFWTLGPFCACTSANNNTYCM 760
Db 719 HRRKKERKEKQKRGEECSLPGLTCFTHDNHWTAPFWNLGSPCACTSSNNNTYCL 778
Qy 761 RTINETHNLFCEFBATGFLYFDLNTDPVQLMNAVNTLDRDLVNLQHLVQLMELRSCGYK 820
Db 779 RTVNETHNFCEFBATGFLYFDLNTDPVQLMNAVNTLDRDLVNLQHLVQLMELRSCGYK 838
Qy 821 QCNPRTRNMDLGLKGGSYEQYQFORRWKPEMKRPSKSLGQWEGWEG 870
Db 839 QCNPRKSLDVGTEGNGYDPR-----GQLWDGWEG 870

```

## RESULT 11

```

Q8C1D3 PRELIMINARY; PRT; 870 AA.
AC Q8C1D3;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Sulfatase FP homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs.";
DR EMBL; AK028285; BAC22858.1; --
SQ SEQUENCE 870 AA; 100903 MW; 12F23402D2F523B5 CRC64;

Query Match 65.8%; Score 3128.5; DB 11; Length 870;
Best Local Similarity 64.9%; Pred. No. 8.3e-240;
Matches 570; Conservative 123; Mismatches 150; Indels 47; Gaps 10;

Qy 5 SLVCLLSATVFSLGSSAFLSHRLKGRFDRRNIRPNILVLTDDQDELGSMQVM 64
Db [1]
4 SLWALLAVLTQLGSLCSTVRSQFRGRIQQRKNIRPNILVLTDDQDELGSLQVM 63
Qy 65 NKTRIMEQGGTHFINAFVTTMCCPSRSSILTKYVHNHNTYTNNECSSPSWQAQHS 124
Db [1]
64 NKTRKIMEQGGATFTNAFTTTPMCCPSRSSMLTKYVHNHNTYTNNECSSPSWQAHEP 123
Qy 125 RTFAVYLNSTGYRTAFPGKYLYNEYSYPPGKWEVGLLKNRFPNYTLCRNGVKEKHG 184
Db [1]
124 RTFAVYLNSTGYRTAFPGKYLYNEYSYPPGKWEVGLLKNRFPNYTLCRNGVKEKHG 183
Qy 185 SDYSKDYLTDLITNDSVSPFRSKMYPHRPVLMVISHAAPHGSDSAPQSRLPFNASQ 244
Db [1]
184 FDYAKDYFTDLITNESINTYFRSKMYPHRPVLMVISHAAPHGSDSAPQSKLYPNASQ 243
Qy 245 HITPSYNYAPNDKHWIMRYTGMPIHMEFTNMLQKRLQTLMSVDDSMETIYNMLVET 304
Db [1]
244 HITPSYNYAPNDKHWIMRYTGMPIHMEFTNMLQKRLQTLMSVDDSMETIYNMLVES 303
Qy 305 GELNDTYIYVTADHGVIHQFGLVKGKSMYPDFIRVPYVRGPNVYAGCLNPHIVLND 364
Db [1]
304 GELNDTYIYVTADHGVIHQFGLVKGKSMYPDFIRVPYVRGPNVYAGCLNPHIVLND 363
Qy 365 LAPTLIDIAGLDIPADMKGKSLKLLDTERPNVRPHLKKQMRWRDSEFLVERGKLLHKRD 424
Db [1]
364 LAPTLIDIAGLDIPADMKGKSLKLLDTERPNVRPHLKKQMRWRDSEFLVERGKLLHKRD 423
Qy 425 NDVKVDAQENFLPKYORVKDLQORABYQACQOLGQKOCVEDATGKLLHKKCKGPMRLG 718
Db [1]
424 ESKGNIQQSNHLPKYERVELCQQAQRYTACQPOQONWOCIEDTSGKLRHKKCKGPMRLG 717
Qy 484 GGSRLSNLVPKYGGGSEACTCDSQDYKLSLAGR---RKKLFKK---KYKASYVRSRSI 537
Db [1]
484 TVHONARNLYSRGLHDKKECHCRDGYRSSRSQRKNQRFLENGTPKYKPRFVHTROT 543
Qy 538 RSVAEIVDGRVHVGLGD-----AAQPRNLTKEHVG-----APEDQDDKGGDFSG 584
Db [1]
544 RLSUVEFEIGEYDINLEEBELQVLPRSSAKRHDEGHQGFIGHQAAAGDIRNMLADSN 603
Qy 585 TGLPDYSAANPIKVTHRCYIILENDVQCDLDLYKSLQAKDKHLHDEIETLQNKIN 644
Db [1]
604 AVGLP-----ATRVTHKCFILPNDTIHCERELYQSARAWKHAYIDKEIEVLQDKIN 658
Qy 645 LREVRGHLKKRPEECCHKISYTHQHG--RLKHRGSSLAHPRK-GLQEKD-KVMLLRE 700
Db [1]
659 LREVRGHLKKRPEECGCGDQSYNKEKGVRQELKSHLPHFKEAAAQEVDSKQLFKE 718
Qy 701 QXKXKLLKLLKXLLQNDTCSMPLGTCFTHDNHWQTAFTWLTGPFCACTSANNTYCM 760
Db [1]
719 HRKKRERKKEKQKRGECBSLGLTCTFDNNHWQTAFTWLTGPFCACTSANNTYCM 778
Qy 761 RTINETHNLFCEFATGFLYFOLNTPDQOLMNAVNTLDRVLNOLHLMELRCKGKY 820
Db [1]
779 RTVNEETHNLFCEFATGFLYFDMNTDPTQLTNVHTVRSILNQLHLMELRSCQGYK 838
```

```
Qy 821 QCNPRTRNMDLGLKDGSEYEQYRQFORRWKPEMKRPSKSLQLWEGWEG 870
Db [1]
839 QCNPRPKSLDIGAKEGGNDPHR-----QQLWDGWEG 870

RESULT 12
Q90XB6 PRELIMINARY; PRT; 867 AA.
AC Q90XB6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE N-acetylglucosamine-6-sulfatase.
OS Coturnix coturnix (Common quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=9091;
RN [1]
RP SEQUENCE FROM N.A.
RA Dhoot G.K.; Gustafsson M.K.; Ai X.; Sun W.; Standiford D.M.;
RA Emerson C.P. Jr.;
RT "Regulation of Wnt Signaling and Embryo Patterning by an Extracellular
RL Sulfatase.";
RL Science 0:0-0(2001).
DR EMBL; AF410802; AAK90515.1; --
DR InterPro; IPR000917; Sulfatase.
DR Pfam; PF00884; Sulfatase; 1.
DR PROSITE; PS00523; SULFATASE_1; 1.
SQ SEQUENCE 867 AA; 100602 MW; 6DA4499EB134A01C CRC64;

Query Match 65.6%; Score 3117.5; DB 13; Length 867;
Best Local Similarity 63.9%; Pred. No. 6.3e-240;
Matches 571; Conservative 117; Mismatches 146; Indels 59; Gaps 10;

Qy 5 SLVCLLSATVFSLGSSAFLSHRLKGRFDRRNIRPNILVLTDDQDELGSMQVM 64
Db [1]
7 ALFLAVLST---ELTSHSSTLKSFLRFRGVRQQRKNIRPNILVLTDDQDELGSLQVM 63
Qy 65 NKTRIMEQGGTHFINAFVTTMCCPSRSSILTKYVHNHNTYTNNECSSPSWQAQHS 124
Db [1]
64 NKTRIMEQGGATFTNAFTTTPMCCPSRSSMLTKYVHNHNTYTNNECSSPSWQAHEP 123
Qy 125 RTFAVYLNSTGYRTAFPGKYLYNEYSYPPGKWEVGLLKNRFPNYTLCRNGVKEKHG 184
Db [1]
124 RTFAVYLNSTGYRTAFPGKYLYNEYSYPPGKWEVGLLKNRFPNYTLCRNGVKEKHG 183
Qy 185 SDYSKDYLTDLITNDSVSPFRSKMYPHRPVLMVISHAAPHGSDSAPQSRLPFNASQ 244
Db [1]
184 FDYAKDYFTDLITNESINTYFRSKMYPHRPVLMVISHAAPHGSDSAPQSKLYPNASQ 243
Qy 245 HITPSYNYAPNDKHWIMRYTGMPIHMEFTNMLQKRLQTLMSVDDSMETIYNMLVET 304
Db [1]
244 HITPSYNYAPNDKHWIMRYTGMPIHMEFTNMLQKRLQTLMSVDDSMETIYNMLVES 303
Qy 305 GELNDTYIYVTADHGVIHQFGLVKGKSMYPDFIRVPYVRGPNVYAGCLNPHIVLND 364
Db [1]
304 GELNDTYIYVTADHGVIHQFGLVKGKSMYPDFIRVPYVRGPNVYAGCLNPHIVLND 363
Qy 365 LAPTLIDIAGLDIPADMKGKSLKLLDTERPNVRPHLKKQMRWRDSEFLVERGKLLHKRD 424
Db [1]
364 LAPTLIDIAGLDIPADMKGKSLKLLDTERPNVRPHLKKQMRWRDSEFLVERGKLLHKRD 423
Qy 425 NDVKVDAQENFLPKYORVKDLQORABYQACQOLGQKOCVEDATGKLLHKKCKGPMRLG 484
Db [1]
424 ESKGNIQQSNHLPKYERVELCQQAQRYTACQPOQONWOCIEDTSGKLRHKKCKGPMRLG 483
Qy 485 GGSRLSNLVPKYGGGSEACTCDSQDYKLSLAGR-----KCLFKKYYKASYVRSRSI 537
Db [1]
484 AIRKRTSRISHRGYSKDKDCGDDTDFNSRTQRKNQRFLENPFSQKYKPRFVHTROT 543
Qy 538 RSVAEIVDGRVHVGLGD-----AAQPRNLTKEHVG-----WFCAPEDQDDKGGDFSG 583
Db [1]
```



```

Db 544 RLSVFEFEIYDINLEEBEOLVLTSTRITKRHNANDKKAEBETGAPGDMVADGTVDI 603
QY 584 GTGGLPDYSAANPIKVTTHRCYILENDTVQCDLDYKSLQAWKHLDHIEITLQNKIK 643
Db 604 G-----QPSSVAVTHKCFILPNDTICERELYSARAKWKHAYIDKSEIALQDKIK 655
QY 644 NUREVRGHLKKRPEECCHKISYTHQHG-----RLKRGSSLHPFRKGLQEKD-KWVL 697
Db 656 NUREVRGHLKKRPPKDECDCTKQSYTNKEKGVKTQEKIK---SHLHPFKEAAQEVDSKLQ 712
QY 698 LREOKRKKLKLKLLKRLONNDTCSMPGLTCFTHDNHWTAPFTWLGPFCACTSSANNNTY 757
Db 713 FENERRRKKRKGKQKKGDECSLPLGTCFTHDNHWTAPFTWLGPFCACTSSANNNTY 772
QY 758 WCMRTINETHNLFCEFATGFLYFDLNTDPYQLMNAVNTLDRVLNQLHVQLMELRSCK 817
Db 773 WCLRTVNDTHNLFCEFATGFLYFDLNTDPYQLMNAVNTLDRVLNQLHVQLMELRSCK 832
QY 818 GYKQCNPRTRNDLGLKDGGSYEQYRQFQRKWPMPKRPSSKSLGOLWEGWEG 870
Db 833 GYKQCNPRPKLETGNKDGGSYDPRH-----GQLMDGWEH 867

RESULT 13
Q8BLJ0 PRELIMINARY; PRT; 862 AA.
AC Q8BLJ0;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Sulfatase Fp homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573 (2002).";
DR EMBL; AK045002; BAC32179.1; --
SQ SEQUENCE 862 AA; 99964 MW; BE67B180D3D26359 CRC64;

Query Match 65.1%; Score 3094.5; DB 11; Length 862;
Best Local Similarity 65.2%; Pred. No. 4.3e-238;
Matches 563; Conservative 121; Mismatches 150; Indels 29; Gaps 9;

QY 5 SLVLCILSATVFSLLGSSAFLSHRLKGRFORDRNRIRPNILVLTDQDVELGSMQVM 64
Db 4 SLWALLAVLGTQLLGLSLCSTVRSQRFGRIOQRKNIRPNILVLTDQDVELGSLQVM 63
QY 65 NKTRIMEQGGTHFINAFVTTMCCPSRSSILTKYVHNHNTYTNENCSSPFWQAQHES 124
Db 64 NKTRKIMEQGGATFTNAFVTTMCCPSRSSMLTKYVHNHNTYTNENCSSPFWQAQHEP 123
QY 125 RTFAVYLNSTGYRTAFEGKLYNEYSYVPPGWKMWGLKXNRFYNTLCRNGVKEKHG 184
Db 124 RTFAVYLNSTGYRTAFEGKLYNEYSYVPPGWKMWGLKXNRFYNTLCRNGVKEKHG 183
QY 185 SDYSKDYLTDLITNDSVSPFRSTKQYRPHRVLWVISHAHPGSDSAPQVSRLPFNASQ 244
Db 184 FDYAKDYFTDLITNESINYPFWSKRMYPHRPIMWVISHAHPGSDSAPQVSKLYPNASQ 243
QY 245 HITPSYNAAPNPKHMYTGPMKDIHMEFTNMLQRLQTLMSVDDSMETIYNMLVET 304
Db 244 HITPSYNAAPNPKHMYTGPMKDIHMEFTNMLQRLQTLMSVDDSMETIYNMLVET 303
QY 305 GELDNTYIVYADHGHYHIGFGLVKGSKMPYEDIRVPFVYRGPVNEAGCLNPHVILNID 364

```

```

Db 304 GELDNTYIVYADHGHYHIGFGLVKGSKMPYEDIRVPFVYRGPVNEAGCLNPHVILNID 363
QY 365 LAPTTILDLAGLDIPADMDGKSLKLLDTEREVNRFHLKQKQNRVWRDSFLVERGKLLHKRD 424
Db 364 LAPTTILDLAGLSDPSVDGKSVLKLDDLEKFGNFRFTNKKAKIWRDTFLVERGKFLAKKE 423
QY 425 NDKVDAQENFLPKYQVVDLCQRAEYQVACEQLQKQWQCVEDATGKLKHKCKGPNRLG 484
Db 424 ESKGNIOQSNHLPKYRVEKELCQQAQYQVACEQGNWQCIEDTSGKLRHKCKGSPDLL 483
QY 485 GSRA-LSNLVPKTYGOGSEACTCDSGDYKLSLAGR---RKCLPKK---KYKASYVRSRSI 537
Db 484 TVRQARNLYSRGLHDKDKECHCRDSGYRSSRSQKQVQFLRNKGTPKYKPRFVHTQT 543
QY 538 RVALTEVDGRVTHVGLGD-----AAQPNMLTTRHWPQ-----APEDQDDKDGDFSG 584
Db 544 RLSVFEFEIYDINLEEBEOLVLTSTRITKRHNANDKKAEBETGAPGDMVADGTVDI 603
QY 585 TGGLPDYSAANPIKVTTHRCYILENDTVQCDLDYKSLQAWKHLDHIEITLQNKIK 644
Db 604 AVGLP-----ATVRVTHKCFILPNDTICERELYSARAKWKHAYIDKSEIALQDKIK 658
QY 645 LREVRGHLKKRPEECCHKISYTHQHG--RLKRGSSLHPFRK-GLQEKD-KWVLRE 700
Db 659 LREVRGHLKKRPEECCHKISYTHQHG--RLKRGSSLHPFRK-GLQEKD-KWVLRE 718
QY 701 OKRKKKRLKRLKRLONNDTCSMPGLTCFTHDNHWTAPFTWLGPFCACTSSANNNTY 760
Db 719 HRRKKRKKRKKRQKGBECSLPLGTCFTHDNHWTAPFTWLGPFCACTSSANNNTY 778
QY 761 RTINETHNLFCEFATGFLYFDLNTDPYQLMNAVNTLDRVLNQLHVQLMELRSCKGYK 820
Db 779 RTVNETHNLFCFATGFLYFDLNTDPYQLMNAVNTLDRVLNQLHVQLMELRSCKGYK 838
QY 821 QCNPRTRNDLGLKDGGSYEQYR 843
Db 839 QCNPRPKSLDIGAKEGGNYDPRH 861

RESULT 14
Q8BX28 PRELIMINARY; PRT; 604 AA.
AC Q8BX28;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Weakly similar to N-acetylglucosamine-6-sulfatase (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573 (2002).";
DR EMBL; AK049170; BAC33584.1; --
FT NON_TER 1
SQ SEQUENCE 604 AA; 69596 MW; A317924B658C59B7 CRC64;

Query Match 64.4%; Score 3059.5; DB 11; Length 604;
Best Local Similarity 92.7%; Pred. No. 1.6e-235;
Matches 561; Conservative 18; Mismatches 19; Indels 7; Gaps 3;

QY 272 HMEFTNMLQRLQTLMSVDDSMETIYNMLVETGELDNTYIVYADHGHYHIGFGLVKGK 331
Db 1 HMEFTNMLQRLQTLMSVDDSMETIYNMLVETGELDNTYIVYADHGHYHIGFGLVKGK 60
QY 332 SMPYEDIRVPFVYRGPVNEAGCLNPHVILNIDLAPTILDIAGLDIPADMDGKSLKLLD 391

```



```
Db 61 SMPYEFDIRVPFVYRGVNEVAGSLNPHIVLNI DLAPTILDIAGLDIPADMDGKSLKLLD 120
QY 392 TERPVNRFHLKKORVWRDQFLVERGKLHKRDNDKVDQAEENFLPKYORVKDLORAEY 451
Db 121 SERPVNRFHLKKORVWRDQFLVERGKLHKRGDKVNAQEENFLPKYORVKDLORAEY 180
QY 452 QTACEQLGQKQWQVEDATGKLKHLKCKGPMRL---GGSRLSNLVPKYTGQSEACTCDS 508
Db 181 QTACEQLGQKQWQVEDASGTLKHLKCKGPMRFGGGGSRLSNLVPKYDQSEACSCDS 240
QY 509 ---GDYKLSLAGRKKLKKKYKASYRSISVAIEVDGRVYHVLGDAAPRNLTGR 565
Db 241 GGGDYKLGLAGRR-KLUFKKKYKTSYARNRSIRSAIEVDGEIYHVGLDTPVPPRNLSKP 299
QY 566 HWPAGEDQDDKDGDFSGTGLPDPYSAANPIKVTHRCYILENDTVQCCLDYKSLQAWK 625
Db 300 HWPAGEDQDDKDGSGTGLPDPYSAANPIKVTHRCYILENDTVQCCLDYKSLQAWK 359
QY 626 DHKLHIDHEIETLQNKIKNREVRGHLKKRPECDCHKISYHTQHGKLRGSSLPFF 685
Db 360 DHKLHIDHEIETLQNKIKNREVRGHLKKRPECDCHKRISYHSQHGKLRGSSLPFF 419
QY 686 RKGLEQDKVWLLREQKKLRKLKRLKRLQNDTCSMPGLTCFTHDNQHWQTAPFTWTLGP 745
Db 420 RKGLEQDKVWLLREQKKLRKLKRLKRLQNDTCSMPGLTCFTHDNQHWQTAPFTWTLGP 479
QY 746 FCACTSANNTYCMRTINETHNPLFCFEPATGFLFYFDLNTDQYQLMNAVNTLDRDLNQ 805
Db 480 FCACTSANNTYCMRTINETHNPLFCFEPATGFLFYFDLNTDQYQLMNAVNTLDRDLNQ 539
QY 806 LHVOLMELRSCGKYQCNPRNMDLGLKDGSGYEVQFORKPKWPKRPSKSLGQLW 865
Db 540 LHVOLMELRSCGKYQCNPRNMDLGLRDSGYSYQFORKPKWPKRPSKSLGQLW 599
QY 866 EGWEG 870
Db 600 EGWEG 604
```

## RESULT 15

```
Q9UPSS PRELIMINARY; PRT; 818 AA.
ID AC Q9UPSS
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein KIAA1077 (Fragment).
GN KIAA1077.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Brain;
EX MEDLINE=93397452; PubMed=10470851;
RA Kikuno R., Nagase T., Ishikawa K., Hirose M., Miyajima N.,
RA Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:197-205(1999).
DR EMBL; AB029000; BAAS3029.1; -.
DR InterPro; IPR000917; Sulfatase.
DR Pfam; PF00884; Sulfatase; 1.
DR PROSITE; PS00523; SULFATASE_1; 1.
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 818 AA; 95067 MW; 43470781E44B871 CRC64;
```

Query Match 64.4%; Score 3059; DB 4; Length 818;  
Best Local Similarity 65.6%; Pred. No. 2.7e-235;

```
Matches 552; Conservative 112; Mismatches 129; Indels 48; Gaps 11;
QY 55 DVELGSMVNNKTRTMEOGGTHFINAFVTTMCCPSRSSILTKYVHNHNTYNNENCS 114
Db 1 DVELGSLQVNNKTRKIMBHGGAFFINAFVTTMCCPSRSSMLTKYVHNHNTYNNENCS 60
QY 115 SPWQAOHESRTFAVYLNSTGYRTAFGKYLNEYNGSYVPPGKWKVGLLKNSREYNTTL 174
Db 61 SPWQAOHESRTFAVYLNSTGYRTAFGKYLNEYNGSYVPPGKWKVGLLKNSREYNTTV 120
QY 175 CRNGVKKHGSYKDYLTDLITNDSVSFFRSKMYPHRPVLAIVISHAAPHGPDSDAPQ 234
Db 121 CRNGIKKHGFDYAKDYFTDLITNESINTYFKMSKRMYPHPPVMMVISHAAPHGPDSDAPQ 180
QY 235 YSRLFNASQHIITPSYNYAPNPKIMRYTCTPMKPIHMEFTNMLQKRLQTLMSVDDSM 294
Db 181 FSKLYPNASQHIITPSYNYAPNPKIMRYTCTPMKPIHMEFTNMLQKRLQTLMSVDDSV 240
QY 295 ETIYNMLETGELDNTYIVYTADHGHIHQPGVLVKGKSMYPYDFIRVPPYVPGNVEAGC 354
Db 241 ERLYNMLETGELDNTYIVYTADHGHIHQPGVLVKGKSMYPYDFIRVPPYVPGNVEAGC 300
QY 355 LNPHIVLNI DLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKOMVWRDQFLV 414
Db 301 IVQIVLNI DLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKOMVWRDQFLV 360
QY 415 ERGKLLHKRDNDKVDQAEENFLPKYORVKDLORAEYOTACEQLGQKQWQVEDATGKLK 474
Db 361 ERGKFLRKKESSKNTQSSNHLPKYERVKELQQARYOTACEQLGQKQWQVEDATGKLRI 420
QY 475 HKCKGPMRLGGS-ALSNLVPKYGGSEACTCDSGYKLSLAGRR-KKLFKK----KY 527
Db 421 HKCKGPMRLGGS-ALSNLVPKYGGSEACTCDSGYKLSLAGRR-KKLFKK----KY 480
QY 528 KASYVRSRSIRSAIEVDGRVYHVLGSD-----AAQPRNLTKRHWPG--APEDQDDKDG 580
Db 481 KPRFVHTQTRSLSVFEGEYIDINLEEEELQVLQPRNIAKRHDEGHGKPRDLQASSGG 540
QY 581 -----DFSGTGGLPQYSAANPIKVTHRCYILENDTVQCCLDYKSLQAWKHLHIDH 633
Db 541 NRGRLADSSNAVGP-----TTVRVTHKCFILPNDISIHCEBELYQSARAWKHKAYIDK 595
QY 634 EIELTQNKIKNREVRGHLKKRPECDCHKISYHTQHGKLRKH--GSSLHPFRKQLQE 691
Db 596 EIELAQDKIKNREVRGHLKKRPECDCHKISYHTQHGKLRKH--GSSLHPFRKQLQE 655
QY 692 KD-KVWLLRE-QIKKKKLRKLKRLQNDTCSMPGLTCFTHDNQHWQTAPFTWTLGPFAC 749
Db 656 VDSKQLQFKENNRKRKRERKQRKGECSLPGTCTFTHDNQHWQTAPFTWTLGPFAC 715
QY 750 TSANNNTYCMRTINETHNPLFCFEPATGFLFYFDLNTDQYQLMNAVNTLDRDLNQHLVQ 809
Db 716 TSANNNTYCMRTINETHNPLFCFEPATGFLFYFDLNTDQYQLMNAVNTLDRDLNQHLVQ 775
QY 810 LMELRSCGKYQCNPRNMDLGLKDGSGYEVQFORKPKWPKRPSKSLGQLWEGE 869
Db 776 LMELRSCGKYQCNPRNMDLGLKDGSGYEVQFORKPKWPKRPSKSLGQLWEGE 817
QY 870 G 870
Db 818 G 818
```

Search completed: October 15, 2003, 13:09:20

Job time : 62 secs